Development of a cellular mechanistic assay for the SET and MYND domain containing methyltransferase SMYD2, identification and validation of a novel substrate, and functional characterization of its inhibition

Dissertation

zum Erwerb des akademischen Grades
Doctor rerum naturalium
(Dr. rer. nat.)
im Fach Biologie

eingereicht an der

Lebenswissenschaftlichen Fakultät
der Humboldt-Universität zu Berlin

vorgelegt von

Erik Eggert

Präsidentin der Humboldt-Universität zu Berlin: Prof. Dr.-Ing. Dr. Sabine Kunst
Dekan der Lebenswissenschaftlichen Fakultät: Prof. Dr. Bernhard Grimm

Gutachter:

1. Prof. Dr. Ana Pombo
2. Dr. habil. Bernard Haendler
3. Prof. Dr. Ingo Morano

Tag der mündlichen Prüfung: 30.06.2017

Contents

List of Figures

List of Tables

Table 0.1: Abbreviations

A	
aa	amino acid(s)
ADP	adenosine diphosphate
ALL	acute lymphoblastic leukemia
ATP	adenosine triphosphate
B	
BCA	bicinchoninic acid
BSA	bovine serum albumin
C	
CCD	charge-coupled device
cDNA	complementary DNA
CDS	coding DNA sequence
CE	crude extract
ChIP	chromatin immuno precipitation
cKO	conditional knockout
CNA	copy number alteration
CRU	central repeated units
CTD	C-terminal domain
D	
DAPI	4',6-Diamidino-2-Phenylindole
DMSO	dimethylsulfoxide
DNA	deoxyribonucleic acid
DRAQ5	1,5-bis{[2-(di-methylamino) ethyl]amino}-4, 8-dihydroxyanthracene-9,10-dione
DTT	dithiothreitol
E	
E	eluate
E. coli	Escherichia coli
e.g.	exempli gratia
EDTA	ethylenediaminetetraacetic acid
ELISA	enzyme Linked Immunosorbent Assay
ESC	embryonic stem cell
ESCC	esophageal squamous cell carcinoma
F	
FCS	fetal calf serum
FDR	false discovery rate
FT	flowthrough
G	
GFP	green fluorescence protein
GST	glutathione S-transferase
H	
hiFCS	heat inactivated FCS
HMT	histone methyltransferase
HRP	horse raddish peroxidase
HS	horse serum
HTS	high throughput screening
I	
IC50	inhibitory concentration 50
ICW	in-cell-western

Table 0.1: Abbreviations

inh.	inhibitor
IP	immunoprecipitation
IPTG	isopropyl -D-1-thiogalactopyranoside
K	
KD	knockdown
KDM	lysine demethylase
KLH	keyhole limpet hemocyanin
KO	knockout
L	
LC-MS/MS	liquid chromatographie – tandem mass spectrometry
LO	lead optimization
M	
M	marker
me	methyl
MOI	multiplicity of infection
MS	mass spectrometry
MW	molecular weight
N	
N.A.	not available
n.d.	not detected
NES	nuclear export signal
NLS	nuclear localization signal
O	
OD	optical density
P	
PAGE	polyacrylamide gel electrophoresis
PBS	phosphate buffered saline
PCR	polymerase chain reaction
PKMT	protein lysine methyltransferase
PMT	protein methyltransferase
PTM	posttranslational modification
Q	
qRT-PCR	quantitative realtime polymerase chain reaction
R	
RNA	ribonucleic acid
RNAi	RNA interference
rpm	rounds per minute
RT	room temperature
S	
S	soluble
SAH	S-adenosyl-L-homocysteine
SAM	S-adenosyl-L-methionine
SDS	sodium dodecyl sulfate
SGC	Structural Genomics Consortium
shRNA	short hairpin RNA
SILAC	stable isotope labeling of amino acids in cell culture
siRNA	short interfering RNA
SPA	scintillation proximity assay

Table 0.1: Abbreviations

T	
TPR	tetratrico peptide repeat
TSA	thermal shift assay
V	
v/v	volume/volume
vol.	volume
W	
W	wash step
w/o	without
WHO	World Health Organization
WT	wildtype
3H-SAM	tritiated SAM

1 Zusammenfassung

Epigenetische Regulatoren gelten als mögliche Angriffspunkte in der Krebstherapie, da sie in vielen Tumoren fehlreguliert sind. Vor allem Protein Methyltransferasen sind gut beschrieben und können durch niedermolekulare Hemmer addressiert werden. Das SET und MYND Domain enthaltene Protein 2 (SMYD2) wurde in der Literatur mehrfach als potenzielles Onkogen beschrieben und eine Überexpression dieses Proteins in Tumorproben war prognostisch für eine schlechtere Überlebensrate und einen aggressiveren Tumor. Für SMYD2 wurden verschiedene Substrate beschrieben u.a. Histon H3 und der Tumor Suppressor p53, allerdings ist die Biologie dieses Enzymes noch kaum verstanden. Durch die Entwicklung einer Testsubstanz zur spezifischen Hemmung des SMYD2 Enzymes könnte ein möglicher therapeutischer Nutzen besser untersucht werden, sowie das allgemeine Verständnis dieser Lysin Methyltransferase vorangetrieben werden.

In der vorliegenden Arbeit sollte daher die Entwicklung eines zellulären mechanistischen Tests zur Messung der enzymatischen Aktivität des SMYD2 Enzyms etabliert werden.
Hierfür wurde ein polyklonaler Antikörper, der gegen ein methyliertes p53 Peptid gerichtet war, generiert. Überraschenderweise detektierte der Antikörper (SY46) nach Überexpression des SMYD2 Enzymes weitere unbekannte Methylierungen im Westernblot. Die zusätzlichen Signale erwiesen sich als spezifisch und wurden für die Etablierung eines mechanistischen Tests genutzt. Mit Hilfe dieses Tests konnte ein spezifischer zellulär aktiver SMYD2 Hemmer identifiziert werden, der aufgrund seiner Selektivität und Potenz für das SMYD2 Enzym als Testsubstanz unter dem Namen BAY-598 der wissenschaftlichen Gemeinschaft zur Verfügung gestellt wurde (für Details siehe Eggert et al.[32]).
Im weiteren Verlauf dieser Arbeit wurde eine Proteomstudie durchgeführt, um zelluläre Lysinmethylierungen zu detektieren, mit dem Ziel neue SMYD2 Substrate zu identifizieren. Dabei wurden hunderte von neuen Lysinmethylierungsstellen entdeckt. Durch Anwendung verschiedener Selektionskriterien wurde das AHNAK Protein als SMYD2 Methylierungssubstrat identifiziert und validiert. Dabei zeigte sich, dass das AHNAK Protein bevorzugt an sogenannten Central Repeated Units (CRUs) innerhalb von LKGPK Sequenzmotiven multi-methyliert werde kann. Die AHNAK Methylierung konnte in verschiedenen Zelllinien und im Muskelgewebe von Mäusen nachgewiesen werden, welches eine konservierte Funktion dieser Modifikation vermuten lässt.
Im letzten Teil der Arbeit wurde die spezifische Testsubstanz BAY-598 genutzt, um exemplarisch verschiedene in der Literatur aufgekommene Hypothesen zur SMYD2 Funktion zu testen. Dabei wurde in einem ersten Ansatz die Differenzierung von Maus C2C12 Myoblasten in muskelähnliche Myotuben unter permanenter SMYD2 Hemmung untersucht. In einem zweiten Zellsystem wurde der RAS Transformationseffekt in NIH3T3 Maus Fibroblasten unter SMYD2 Hemmung oder Knockout getestet, da es in vivo Hinweise gab, dass SMYD2 die RAS getriebene Adenocarcinomentstehung im Pankreas beeinflusst.

Die vorliegende Arbeit hat mit dazu beigetragen die potente und selektive SMYD2 Testsubstanz BAY-598 zu entwickeln. Außerdem wurde mit AHNAK ein neues SMYD2 Substrat identifiziert und validiert. Diese Methylierung ist nicht nur in humanen Zellkulturen nachweisbar sondern auch im Muskelgewebe der Maus. Die Relevanz des SMYD2 Enzymes und der AHNAK Methylierung erfordert weitere Forschungsarbeit, die durch die Bereitstellung der spezifischen Testsubstanz BAY-598 deutlich verbessert werden sollte.

2 Introduction

2.1 Cancer is an emerging burden to the society

Cancer is after heart disease the second most common cause of death in western countries. Due to the aging population, it is estimated that the worldwide cancer incidence will increase[124]. According to the WHO World Cancer Report 2014[124], there were 14 million new cases and 8.2 million cancer-related deaths worldwide in 2012. The number of new cancer cases is expected to increase to 22 million within the next two decades. In the United States, the lifetime risk of developing cancer on average is about 1 in 2 in men and about 1 in 3 in women[56]. Although in the US the 5-year relative survival rate for all cancers has improved from about 49% in the 1970's to about 69% in the 2010's basically due to better diagnosis and improved treatment, there is still a high unmet medical need for better treatment options of cancer patients.
Cancer is a generic term for a large group of malignant neoplasms, which can affect any tissue of the body. The most common cases of cancer related death worldwide are malignancies of lung (1.59 million deaths), liver (745 000 deaths), stomach (723 000 deaths), colorectal (694 000 deaths), breast (521 000 deaths), and esophageal cancer (400 000 deaths)[124]. One defining feature of cancer is the abnormal cell proliferation beyond their usual boundaries and the ability to invade surrounding tissue and spread to other parts of the body, which is known as metastasizing. Metastases are the major cause of death from cancer.
With a better understanding of the molecular mechanisms that drive oncogenesis, it has become apparent that tumors originating from the same organs can display very different molecular profiles and underlying mutations[134]. This dramatically influences prognosis and response to treatment and requires different therapeutic intervention. Targeted therapies have emerged as a promising strategy to more specifically tackle cancer cells based on molecular categorization and may avoid unintended side effects that arise from unspecific chemotherapy[9]. For instance, specific inhibition of EGFR in tumors that display hyperactivation of this tyrosine kinase receptor significantly improved the outcome in this patient population. However, targeted therapy turned out to be more challenging as initially anticipated due to a lack in understanding the complex underlying biology and pathobiology and more research is needed to identify effective targets in non-responding patients.

For targeted therapy in cancer, potential vulnerabilities have been hypothesized and in 2000 Hanahan and Weinberg defined the so called "hallmarks of cancer" that are mandatory for normal cells to acquire in order to develop a full-blown cancer[48]. These hallmarks were defined as sustained proliferative signaling, evasion of growth suppressors, resistance to cell death, induction of angiogenesis, replicative immortality, and activation of invasion and metastasis. A decade later, the list of original hallmarks was extended by deregulation of cellular energetics, genome instability and mutations, avoidance of immuno destruction, and tumor promoting inflammation (**Fig. 2.1**)[49].

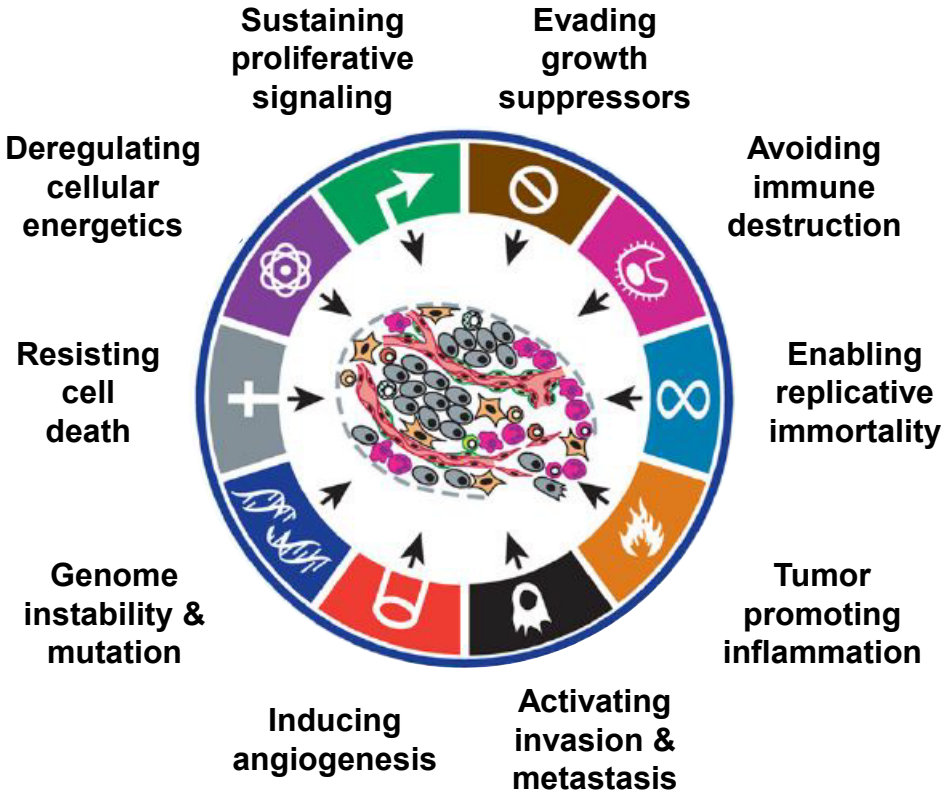


Fig. 2.1: Hallmarks of cancer.
Cartoon representing the updated "hallmarks of cancer" as defined by Hanahan and Weinberg. The cartoon was modified from the corresponding publication [49].

Originally, it was thought that these hallmarks were solely acquired by genetic mutations. However, more recent data suggest that many of the changes during transformation are caused by epigenetic alterations, which act in concert with genetic mutations [139]. For instance, in many tumors DNA methylation patterns, histone modifications and nucleosome positioning differ dramatically between healthy and tumoral tissue, which causes important changes in gene expression and behavior of the cells. Because epigenetic changes can occur without changes in the underlying DNA sequence, it is believed that these changes are reversible and may allow a correction of the abnormal cancer phenotype by specific interventions. Therefore, epigenetic modifiers have emerged as promising new targets. Indeed, the first generation of DNA methyltransferase inhibitors is already on the market to treat myelodysplastic syndromes [101] and more epigenetic target classes are considered for cancer therapy in preclinical and clinical research. Among the diverse classes, protein methyltransferases (PMTs) are well-defined enzymes with reasonable drugability and are often misregulated in cancer and other diseases [51]. In fact, inhibitors for two protein methyltransferases, namely DOT1L and EZH2 have already entered clinical trials and more are expected to come [88].

2.2 Protein methylation & Protein methyltransferases (PMTs)

Protein methylation was first discovered in 1959 in a bacterial flagellar protein[5] and a few years later, this posttranslational modification (PTM) was also found on mammalian histones[89]. In general, protein methylation is enabled by the transfer of a methyl group (–CH3) from the cofactor S-adenosyl-L-methionine (SAM; also known as AdoMet) to several amino acid residues of target proteins with the byproduct S-adenosyl-L-homocysteine (SAH)[20]. SAH can then be hydrolyzed to homocysteine and adenosine by S-adenosyl-L-homocysteine hydrolase. In a second step, methionine synthases catalyze the transfer of a methyl group from 5-methyltetrahydrofolate to homocysteine to produce methionine. SAM is finally recycled from methionine and ATP by methionine adenosyltransferase[37]. It is estimated that SAM is the second most widely used enzyme co-factor after ATP indicating that methylation in general (of DNA, RNA, proteins, metabolits, etc.) plays important roles within the cells. However, the metabolic price for methylation is high. Whereas a kinase consumes 1 ATP equivalent for catalyzing phosphorylation, methylation reactions using SAM as a methyl donor comes with a cost of 12 ATP equivalents, thus active methylation is one of the most expensive reactions per carbon basis[7]. Having said that, the biological relevance of individual methylation reactions might be put under constant evolutionary pressure.

Enzymes responsible for protein methylation reactions can be classified into distinct protein classes based on the target amino acid of the methylation reaction. In human, two classes of PMTs have been identified and are referred to as protein arginine and protein lysine methyltransferases (PRMTs and PKMTs), respectively[12]. PRMTs catalyze the transfer of a methyl group from SAM to the guanidino nitrogen atoms of arginine residues, which can result in arginine mono- or dimethylation, whereby dimethylation can occur either symmetrically (one methyl group per terminal nitrogen atom) or asymmetrically (two methyl groups on a single terminal nitrogen atom)[10]. The other important class comprises the protein lysine methyltransferases (PKMTs), which catalyze the methylation of the ϵ-amino group of lysine residues. This can result in mono-, di-, or trimethylation (**Fig. 2.2**)[75]. The addition of multiple methyl groups to lysine residues might be catalyzed by a single PKMT or might be catalyzed subsequently by different enzymes. For instance, NSD family members can mono- and dimethylate H3 K36, the latter one of which serves as the substrate for SETD2-dependent trimethylation[73;132]. In contrast, EZH2 is able to generate mono-, di-, and trimethylated lysines on H3 K27 albeit with different catalytic efficiencies[121]. Proteins can be methylated at different lysine residues by different PKMTs and single lysine residues on certain proteins can be the target of unique PKMTs or may be a target of many PKMTs. For instance, H3 K79 is a unique target of DOTL1 methylation[92], whereas H3 K9 has been reported to be methylated by at least eight PKMTs[136]. On the other hand, one PKMT can have many substrates[26]. Based on initial experiments showing a half-life of histone methylation similar to the half-life of histones themselves, it was suggested to be an irreversible posttranslational modification (PTM)[8]. However, due to the discovery of lysine-specific demethylase 1 (LSD1; also known as KDM1A)[116], as well as the Jumonji C (JmjC)-domain-containing protein family[130], both of which could be shown to possess protein demethylase activity, it became apparent that protein lysine methylation is dynamically regulated.
Initially, PMTs were most commonly referred to as histone methyltransferases (HMTs), because most discoveries and characterizations were done on histones. However, more recent discoveries are indicating that histones constitute only a subset of all potential protein tar-

gets within the cell[54]. With a few exceptions like DOTL1[35] and METTL21A[117], which are structurally closer to PRMTs, most PKMTs contain a SET domain, which is responsible for the catalytic activity. The SET domain is approximately 130 amino acids in length and was originally identified in the Su(var)3–9, Enhancer-of-zeste and Trithorax proteins, from where the name was derived[58;129]. It is defined by specific amino acid motifs (ELxF/YDY and RFINHxCxPN, where x is any amino acid) and a pseudoknot structure[109]. In addition, all SET domain containing PKMTs also possess I-SET (immunoglobulin-SET) and post-SET domains, which both contribute to peptide and SAM substrate recognition[109].

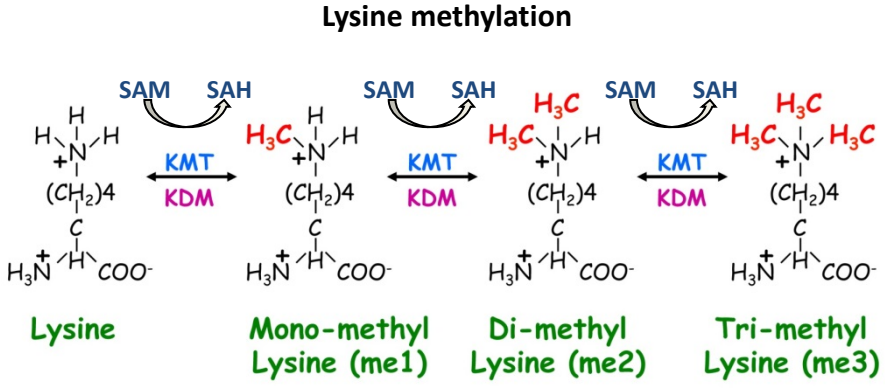


Fig. 2.2: Lysine methylation.
Representation of the amino acid lysine and subsequent methylation of the ϵ-amino group from non-methylated to tri-methylated. Methylation reactions are catalyzed by lysine methyltransferases (KMTs) using cofactor S-adenosyl-L-methionine (SAM) as methyl donor, which is converted into S-adenosyl-L-homocystein (SAH). KMTs are either specific for methylation reactions of specific methylation states or can catalyze all three methylation states. Methylated lysines can be demethylated by lysine demethylases (KDM) such as LSD1[104] indicating a dynamic regulation within cells. Reaction scheme was adjusted from website of Gozani lab (`http://web.stanford.edu/group/gozani/cgi-bin/gozanilab/research`)

2.3 PKMTs - the SMYD family

Currently, more than 50 SET domain containing PKMTs have been identified in the human genome and based on sequence similarity of the functional domains these can be divided into several families (**Fig. 2.3**)[6]. Among them, the SMYD gene family is defined by a conserved core comprising a SET domain interrupted by a MYND zinc finger domain. The mammalian genomes encode for five SMYD genes named SMYD1-5. Crystal structures are available for human SMYD1-3. Their overall structure and domain arrangement is very similar with an N-terminal SET and MYND domain followed by a C-terminal domain (CTD), which structurally resembles tetra-trico peptide repeats (TPR) and may be involved in protein-protein interactions. Overall, they share sequence identity of approximately 31% among each other, which is mainly due to the conserved SET and MYND domains. SMYD4 has an additional TPR-like domain at the N-terminus in front of the SET and MYND domain, whereas SMYD5 has no TPR-like domain. Whereas orthologs of SMYD3, SMYD4 and SMYD5 can be found already in early metazoans, Calpena et al. suggested that chordate-specific SMYD1 and vertebrate-specific SMYD2 evolved from gene duplication events of an ancestral gene of SMYD3[15].

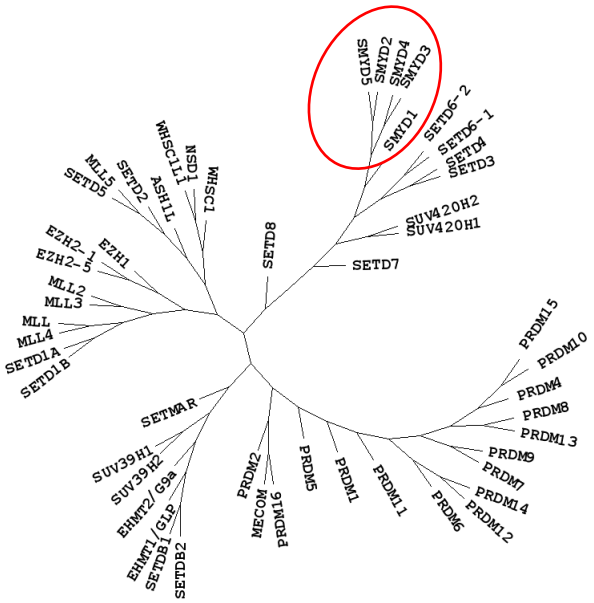


Fig. 2.3: Phylogenetic tree of human SET domain containing PKMTs.
Phylogenetic tree of human SET domain containing PKMTs based on sequence homology of the catalytic SET domain. The SMYD protein family is highlighted in red. The phylogenetic tree was derived from the ChromoHub of the SGC homepage [81].

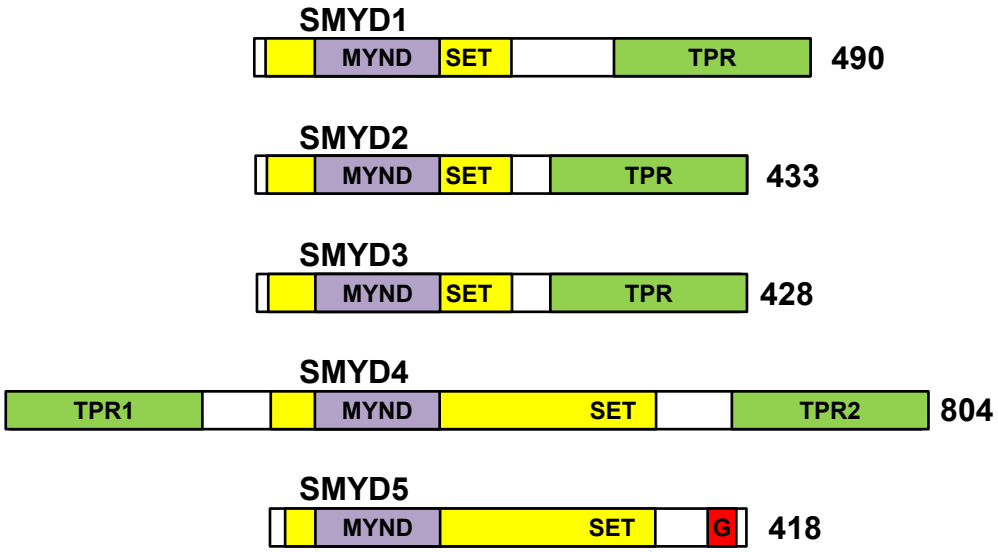


Fig. 2.4: Domain structure of human SMYD protein family.
The human SMYD family comprises five members named SMYD1-5. They are characterized by a MYND domain, which interrupts the catalytic SET domain. SMYD1-3 have an additional C-terminal TPR like domain, whereas SMYD4 comprises one N-terminal and a second C-terminal TPR like domain. SMYD5 lacks TPR like domains and has a glutamine rich C-terminus. The numbers at the end indicate the number of amino acids of the polypeptides. The cartoon from the domain architecture was adjusted from Abu-Farha et al. [2]

Generally, SMYD proteins have been assigned functional roles in cardiomyogenesis and myofibrillogenesis[31] as well as in oncogenesis[45]. The most significant function has been demonstrated so far for SMYD1 also known as BOP, which is mainly expressed in cardiac and skeletal muscle cells. Homozygous knockout (KO) of SMYD1 in mice resulted in embryonic lethality by disrupting maturation of ventricular cardiomyocytes[41]. Similar loss of function studies in zebrafish also indicated an important role for SMYD1 in myofibril organization and muscle contraction[127]. Additional studies in zebrafish suggested that SMYD1 is localized to the sarcomeric M-line, where it physically associates with myosin[61]. In vitro methylation assays suggested methylation activity towards histones[127]. However, localization of SMYD1 is mainly in the cytoplasm[79].
SMYD3 is mostly cited in the context of cancer as a potential oncogene[17;44;46] and very little is known about its normal physiological function. It is more broadly expressed compared to SMYD1, but similar to SMYD1 it has been linked to cardiac and skeletal muscle function, as SMYD3 specific antisense morpholino-oligonucleotides in zebrafish resulted in disturbed muscle development and inappropriate expression of heart-chamber markers and myogenic regulatory factors[39]. However, in contrast to SMYD1, dependency of SMYD3 in zebrafish muscle development was not observed in mice, as SMYD3 knockout mice were viable, fertile, and showed no obvious phenotype[86]. The same group who generated complete SMYD3 knockout mice also generated pancreas specific knockouts and showed that RAS induced tumor formation in pancreas was retarded in SMYD3-KO mice. They discovered MAP3K2, a mitogen activated kinase family member, as a new substrate of SMYD3. Methylation of MAP3K2 may prevent its dephosphorylation at neighboring residues, which led to extended kinase activity downstream of RAS signaling. This might have promoted tumor formation.

Very little is known about SMYD4 and SMYD5. A homolog of SMYD4 in fruit fly was reported to be a muscle-specific transcriptional modulator involved in development[128]. A second study in cellular breast cancer models suggested a potential tumor suppressor function for SMYD4 by suppressing the expression of platelet derived growth factor receptor alpha (PDGFRα)[53]. SMYD5 was suggested to methylate H4 K20 based on biochemical methylation assays on recombinant histones. This was associated with negative regulation of inflammatory response genes[123]. A very recent report described SMYD5 as an important player in hematopoiesis during zebrafish embryogenesis[38]. However, more work is needed to better understand the functions of SMYD4 and SMYD5.

2.4 SET and MYND domain containing 2 - SMYD2

SET and MYND domain containing 2 (SMYD2), also known as KMT3C, HSKM-B, or ZMYND14 is a SET domain containing protein lysine methyltransferase. The human SMYD2 gene is encoded on chromosome 1 at 1q32.3 and the mature transcript (NM_020197.2) derives from 12 exons. The coding sequence (CDS) encodes for a protein of 433 aa (NP_064582.2) with a predicted molecular weight of 49.5 kDa. No alternative splicing variants or protein isoforms have been described so far.

SMYD2 is a multi-domain protein and consists primarily of two lobes, that are separated by a deep groove (**Fig. 2.5**)[36]. The N-terminal lobe (residues 1-279) comprises the catalytic SET domain, the MYND domain, the insertion SET-I domain, and the post-SET domain. The catalytic SET domain within the N-terminal lobe is split into an N-terminal S-sequence (residues 1–49) and a core SET domain (residues 183–246). In between lies the

MYND domain (for Myeloid, Nervy, and DEAF-1, residues 50–98). Whereas the SET domain is responsible for the methyltransferase activity towards lysine residues[28], the MYND domain with its conserved zinc-finger motif can mediate protein-protein interactions by binding to proline-rich sequences[122]. For instance, it was shown that SMYD2 binds to proteins such as the cytoskeleton associated protein EBP41L3 via interaction of its MYND domain with PxLxP motifs[1]. Though interrupted in its primary sequence, the catalytic SET domain occupies a similar overall fold compared to other SET domain containing proteins[59]. SET-I domain (residues 101–182) and the cysteine rich post-SET domain (residue 246–282) are thought to support SAM cofactor and peptide substrate binding[21]. Three highly conserved sequence motifs can be found within the catalytic SET domain of SMYD2 and are essential for efficient SAM binding[18]. These are the GxG motif (residues 18–20 of the S-sequence), the NHxCxPN motif (residues 206–212 of the core SET domain), and the GEExxxxY motif (residues 233–240 of the core SET domain). Disruption of these motifs have been shown to interrupt efficient SAM binding and therefore methyltransferase activity[1;137]. The C-terminal lobe (residues 280-433) comprises one domain, referred to as C-terminal domain (CTD) or tetratrico peptide repeat (TRP) like domain[120], which is less well characterized. The CTD spans around 150 aa among most SMYD-family members. It is composed of seven antiparallel α-helices with an overall structural fold similar to tetratricopeptide repeats (TPRs). Therefore, the CTD has been suggested to be involved in protein-protein interactions[59].

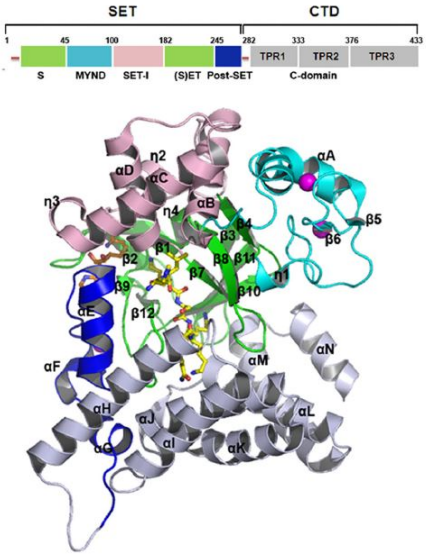


Fig. 2.5: Human SMYD2 protein structure.
Domain representation of SMYD2 (top). The domains are indicated by color and numbers indicate amino acid residues that define the boundaries of the domains. Bottom, shows a ribbon diagram of SMYD2 structure. Individual domains are colored according to the above domain scheme. The co-factor SAH is shown as a brown stick model, a p53 substrate peptide is shown as a yellow stick-and-ball model, zinc ions are shown as magenta spheres. The figure was taken from Wang et al.[133].

2.4.1 SMYD2 in normal physiology

Several groups have tried to uncover the function of SMYD2 in normal physiology. In 2006 Brown et al. were the first, who published a functional characterization of SMYD2[13]. In that report, the enzyme had been described to methylate H3 K36 based on in vitro methyltransferse assays using recombinant histones and methylation specific antibodies. Furthermore, the authors suggested an inhibitory activity of SMYD2 towards transcription, based on results from a luciferase gene reporter assay using SMYD2-GAL4 fusion protein and an SV40 promotor with GAL4 binding sites. The inhibitory activity was explained by observed interactions of ectopic SMYD2-GAL4 with ectopic histone deacetylase HDAC1 and endogenous Sin3A upon immunoprecipitation, the last two, which have been shown to form transcriptional repressor complexes[66]. By northern blot analyses they detected high transcript levels of SMYD2 in various mice tissues, especially in the brain, the heart and the developing embryo. Kawamura et al. confirmed high expression of SMYD2 in cardiac and skeletal muscle in *Xenopus laevis*[65]. This was further confirmed by Diehl et al. on RNA level and protein level in tissues from mice and rat. This group not only showed a high expression of SMYD2 in the postnatal heart, but they were also the first who published data from a genetic knockout of SMYD2 in mice[27]. They generated conditional knockout (cKO) mice harboring a cardiomyocyte-specific deletion of SMYD2. Unexpectedly, in contrast to its high expression in the neonatal heart and its regulation during heart development, SMYD2-KO mice were functional and morphological indistinguishable from control mice, suggesting no non-redundant contribution to the development and maintenance of normal heart physiology. Importantly, global H3 K4 and K36 methylation was not affected in SMYD2-KO cells, challenging the relevance of initial findings of SMYD2 activity towards these histone marks. Donlin et al. discovered SMYD2-dependent lysine methylation of HSP90 at residue K616 (K615 in human) in several cell lines including muscle cells, which enabled complex formation between SMYD2, HSP90 and the sarcomeric protein titin[30]. Specifically, methylation of HSP90 was shown to be necessary to enable complex formation of SMYD2 and HSP90 with the N2A domain of titin at the I-band of the sarcomeres. A subsequent study by Voelkel et al. confirmed SMYD2-mediated complex formation of methylated HSP90 and the N2A domain of titin[131]. Interestingly, in contrast to mammals, downregulation of SMYD2 in zebrafish was reported to disrupt the stability of the sarcomeric titin springs, which resulted in severely impaired mobility, malformation of the tails, reduced heart rates, and fractional shortening[131]. This suggested a compensatory mechanism for SMYD2 in rodents. SMYD2 expression has also been shown to be induced during human embryonic stem cell (ESC) differentiation[113]. Connecting findings from both cancer research and cardiac muscle physiology, Sajjad et al. investigated the p53-dependent apoptotic induction in cardiomyocytes upon hypoxia in the context of SMYD2 activity[106](p53 was described as a SMYD2 substrate as discussed later). They found that SMYD2 is downregulated upon apoptotic triggers both on the transcriptional and post-translational level. They further suggested that SMYD2 had an inhibitory effect on apoptosis induction by the observation that SMYD2-null cardiomyocytes revealed a stronger apoptotic response upon $CoCl_2$-induced hypoxia as an apoptotic trigger. However, the strong expression of SMYD2 during development and in cardiac and skeletal muscle remains a mystery in view of the lack of phenotype in SMYD2 knockout mice (besides cardic-specific SMYD2-KO mice, another group has generated complete knockouts with no observed phenotype[103]). Abu-Farha and colleagues combined proteomic and genomic approaches to identify the SMYD2 interactome and its impact on gene regulation[1]. Potential protein interaction partners where identified by mass spectrometry

after immunoprecipitation of ectopic FLAG-tagged SMYD2 using the human embryonic kidney derived 239T cell line. The authors identified 21 potential SMYD2 interacting proteins. Further validation experiments by co-immunoprecipitation of ectopic SMYD2 variants and potential interactors revealed interactions of SMYD2 with HSP90 independent of its SET and MYND domains, with p53 dependent on functional SET domain and of EPB41L3 dependent on the MYND domain. The interaction of EPB41L3, a membrane and cytoskeletal associated protein was presumed to depend on a PxLxP motif within the protein. Furthermore, they showed that the interaction of HSP90 with SMYD2 enabled SMYD2 to mono-methylate H3 K4 in biochemical methylation assays. They also observed mainly upregulation of genes upon ectopic SMYD2 overexpression in 293T cells. They hypothesized that upregulation was due to the ability of SMYD2 to methylate H3 K4 on the corresponding promotor regions.

2.4.2 SMYD2 function in cancer

Huang et al. were the first who published the ability of SMYD2 to methylate the tumor suppressor p53 and thereby inhibiting p53 transactivation activity[55]. Specifically, the authors were able to demonstrate, both in biochemical and in cellular assays, that K370 of p53 is a specific target lysine for monomethylation by SMYD2. They further showed that modulation of SMYD2 abundance by either knockdown or overexpression, respectively increased or decreased the p53-dependent induction of canonical p53 target genes p21 and mdm2. In addition, knockdown of SMYD2 also resulted in a pronounced p53 mediated apoptosis induction in the U2OS osteosarcoma cell line model (p53 WT) upon treatment of the DNA damage-inducing agent Adriamycin. Using a double chromatin immunoprecipitation approach (ChIP 1 against total p53, ChIP 2 against methylated p53) the authors then provided evidence for a mechanistic model, by which p53 K370 mono-methylation reduces the ability of p53 to bind chromatin and thereby impairing transcriptional induction of p53 target genes. Importantly, these findings indicated a potential oncogenic activity for SMYD2. Subsequently, several crystal structures were published of SMYD2 in conjunction with p53 derived peptides, which supported p53 regulation by SMYD2 methylation[36;133]. In 2009, Komatsu et al. published that SMYD2 lies in a frequently amplified chromosomal region within 1q32–q41 in esophageal squamous cell carcinoma (ESCC) patients and that overexpression of SMYD2 in ESCC correlated with a poor prognosis of these patients[70]. What is more, this correlation was considered to be not only valuable as prognostic marker but also a causal driver of oncogenesis, as knockdown of SMYD2 expression inhibited and ectopic overexpression of SMYD2 promoted the proliferation of ESCC cell lines. However, SMYD2 overexpression was independent of the p53 mutation status and prompted towards additional roles of SMYD2 besides p53 regulation. Similar results were later found in gastric cancer[69], in bladder carcinomas[19], in acute lymphoblastic leukemia (ALL)[107] and in human papillomavirus (HPV)–unrelated head and neck squamous cell carcinomas (HNSCC)[95]. All these reports pointed towards a critical role of SMYD2 in tumorigenesis and a potential target for cancer therapy. Mechanistically, SMYD2 methylation activity has been linked to several substrates besides p53 and histone H3. For instance, it has been reported that retinoblastoma protein (RB), a tumor suppressor and a major regulator of cell cycle progression at the G1/S phase, can be methylated by SMYD2 at residue K860[105]. The authors also detected an increase in RB methylation upon DNA damage caused by Etoposide treatment. In this study, it was furthermore indicated that monomethylation of RB at K860 may serve as a binding motif for the L3MBTL1 transcriptional repressor via its MBT domain. This hypothesis was based on co-immunoprecipitation experiments of

ectopic RB and ectopic L3MBTL1 in the presence or absence of ectopic SMYD2. Overall, these data suggested a repressive function of RB methylation towards RB target genes and an antiproliferative effect of SMYD2 activity. This is in contrast to another study, where Cho et al. discovered SMYD2-dependent RB methylation at K810, which promoted cell cycle progression by increasing neighboring RB phosphorylation at S807/811[19]. Additional studies discovered the nuclear hormone receptor estrogen receptor alpha (ERα) to be a direct target for protein methylation at residue K266 by SMYD2[141]. Methylation at K266 reduced the ability of ERα to bind to its target genes. However, stimulation with estradiol diminished ERα methylation and subsequent target gene activation. Later on, a crystal structure was published of SMYD2 with a peptide ligand derived from ERα containing the methylated target lysine K266[60]. Hamamoto et al. showed that FLAG-tagged SMYD2 can immunoprecipitate HSP90AB1 and is able to monomethylate HSP90AB1 at residues K531 and K574, which increased its dimerization ability and supported cancer cell growth. Interaction of SMYD2 with HSP90 was shown to be dependent on the SET domain of SMYD2 and the C-terminal region of HSP90[47]. This is in contrast to Abu-Farha et al. who indicated an interaction of SMYD2 with HSP90 independent of the SET and MYND domain. Piao et al. reported on findings showing that SMYD2 is able to methylate Poly(ADP-ribose) polymerase-1 (PARP1) at residue K528[98]. PARP1 is highly abundant in the nucleus and involved in DNA repair, chromatin modification, transcriptional regulation and genomic stability[62;94]. Methylation of PARP1 increased its poly-(ADPribosyl)ation (PAR) activity in biochemical assays. Furthermore, knockdown of SMYD2 decreased PAR activity in cells after DNA damage and oxidative stress as measured in immuno-fluorescence analyses by a PAR-specific antibody. Nakakido et al. reported that SMYD2 methylates PTEN at residue K313 in vitro and in vivo[91]. The authors suggested a model by which methylation of PTEN at K331 supports PTEN phosphorylation at S380 and thereby inactivating PTEN tumor suppressor function and activation of the PI3K – AKT signaling pathway. In contrast, knockdown of SMYD2 reduced AKT activity, as measured by AKT phosphorylation on residue T308, and resulted in growth inhibition of breast cancer cell lines. Taken together, multiple regulatory functions have been assigned to SMYD2 towards prominent oncogenes or tumor suppressor genes. In addition, it is frequently overexpressed in many tumor types and is associated with a worse outcome. Therefore, SMYD2 may represent a promising anti-cancer target. The fact that SMYD2 knockout mice are viable and show no phenotype may predict good tolerability of healthy tissue upon SMYD2 inhibition and therefore may allow a targeted therapy specific to cancer cells. This is the reason, why pharmaceutical companies started drug discovery projects for the development of potent small molecule SMYD2 inhibitors.

2.5 Drug discovery and development

For the development of targeted therapies highly selective and potent small molecules or biologics are needed. This requires a comprehensive drug discovery and development process, which is cost and time consuming, often with a time frame from "first hit" to market launch of over ten years and costs exceeding on average $ 1 billion[29]. The prerequisite to engage in a drug discovery program are defined first by an unmet clinical need and second by a scientific hypothesis, often originating from basic science, which suggests a potential benefit for patients when targeting a certain protein or pathway in the underlying disease. The preclinical stages of drug discovery start with initial target identification (either from phenotypical screenings or from literature mining) and to some extent target validation. However, the latter one is only partially feasible due to technical limitations and time.

Subsequently, assays need to be developed that are suitable for high-throughput screening (HTS) in order to discover suitable chemical structures as starting points for further optimization towards the defined target. The HTS will identify initial hits from a compound or biological library. Upon confirmation of initial hits, certain leads with most promising properties will be selected for further lead optimization (LO). Parallel to the LO phase, which is driven in the case of small molecule inhibitors mainly by medicinal chemistry, more comprehensive in vitro and in vivo pharmacological characterization is performed to test for cellular activity of the generated compounds. For that purpose, cellular mechanistic assays are highly desired that adequately monitor the specific activity of a target. At the end of the LO phase, best candidates are chosen for comprehensive toxicology and pharmacokinetic/pharmacodynamic testing in animal models before the "first in man" trials can be initiated in phase 1 clinical studies. This marks the transition from preclinical to clinical development.
The clinical development normally consists of phase 1-3 clinical studies after which approval from healthcare authorities may be given, if effectiveness and safety of the investigational new drug can be demonstrated at a proper risk/benefit ratio[3]. However, there is a big attrition on the way to market launch and advanced drug candidates can fail for two main reasons: Either they are not effective or they are not safe. Lack of efficacy and safety can be due to the under- or overestimated consequences of inhibiting or activating a certain target of interest or due to unintended off-target effects. Although many of the effects are difficult to predict, especially when transferring from preclinical models to human, a better understanding of the biology of individual targets can significantly mitigate late stage failure.

High quality small molecule probes can make a contribution to generate a more efficient drug discovery and development process, because they will allow a broad scientific community to work with these tools and advance our understanding of proposed targets and may even identify new unexpected applications or indications for certain targets. Small molecule probes are especially valuable as they are easy to handle (e.g. treating cells with a compound is much easier accomplished than transfecting them with siRNAs or creating knockouts), possess completely different dynamics compared to genetic tools like RNAi mediated knockdown, and often better mimic the pharmacological properties of drugs in the clinic. In addition, phenotypes can greatly vary between the depletion of a protein by knockout/knockdown or the inhibition of a specific activity of that protein (e.g. enzymatic activity, protein-protein interaction, protein-DNA interaction), which otherwise may compromise interpretation of target characterization. Therefore, high quality potent and selective drug-like probes that are suitable for in vitro and in vivo research are highly desired by both academic research and pharmaceutical industry to enhance the scientific progress and enable faster delivery of better drugs to patients.

2.6 Previous research activities as starting point for the current thesis

Due to a high need for better anti-cancer therapies and the hints coming from research that SMYD2 might be involved in oncogenesis, this enzyme was suggested as a therapeutic target. However, the biology of SMYD2 remains elusive and there is still an ongoing debate about its relevant substrates and molecular mode of action during tumorigenesis and further exploration is needed to solidify potential therapeutic benefits for patients. Therefore, it was envisioned to develop a specific in vivo probe for SMYD2 to allow more research on the impact of pharmacological inhibition of SMYD2 in diverse model systems.

For that purpose a biochemical scintillation proximity (SPA) assay was established, which quantitatively measured the methylation activity of recombinant human SMYD2 enzyme towards a p53 substrate peptide. The assay was used for screening a nearly 3 Mio compound library to identify chemical structures that were able to inhibit SMYD2 activity. Initial hits were reconfirmed and tested in an orthogonal thermal shift (TSA) assay. Finally, an aminopyrazoline molecule was selected based on physico-chemical properties for further chemical optimization (**Fig. 2.6**) (For details on the screening procedure and on individual assays see Eggert et al.[32]). As mentioned before, during the process of compound optimization, a mechanistic assay was needed to specifically determine the potency of individual compounds on cellular SMYD2 activity.

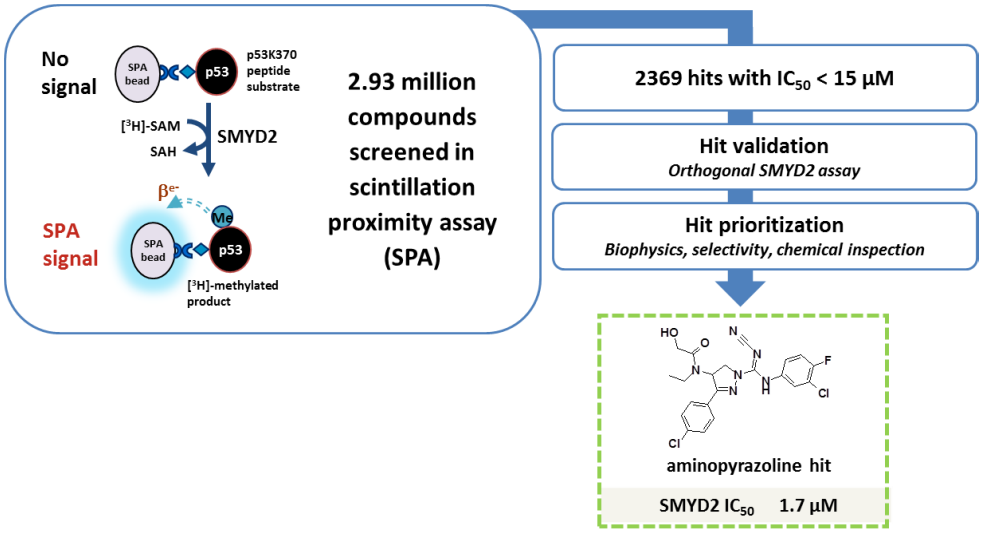


Fig. 2.6: Scheme of SMYD2 inhibitor screening campaign. To identify suitable chemical structures for the development of SMYD2 inhibitors a biochemical scintilliation proximity (SPA) assay was used with biotinylated p53-peptide containing lysine residue K370, recombinant SMYD2 enzyme, and radio-labeled SAM. Upon transfer of tritiated methylgroup from SAM to p53 peptide, β-emmission in proximity to the SPA beads releases photons from the SPA beads that can be quantified. The assay was used to screen a nearly 3 Mio. compound library. Initial hits were prioritized after orthogonal assay validation and selection towards suitable physico-chemical properties. A aminopyrazoline based chemical structure showed promising initial potency with IC50 of 1.7 µM and was selected for further optimization. For details see Eggert et al.[32].

2.7 Aim of this thesis

To continue in the development of a SMYD2 probe after screening, similar to the above described process for drug discovery and development, it was my first task to support the establishment of a SMYD2 specific cellular mechanistic assay that is able to monitor the enzymatic activity and the effects of small molecule inhibition.
Secondly, I intended to discover novel substrates of SMYD2 that might enable a better understanding of that enzyme in the cellular context.
Finally, I aimed at testing the biological relevance of SMYD2 activity in selected cellular models by genetic depletion of SMYD2 and by inhibiting SMYD2 activity with BAY-598, a potent and selective SMYD2 probe that was developed in the course of this project. The latter one should be seen as a starting point towards additional research about SMYD2 protein and the suitability of using probes to address specific scientific questions.

3 Materials & Methods

3.1 Mammalian cell culture

All cell lines were purchased from the American Type Culture Collection (ATCC) or the German Collection of Microorganisms and Cell Cultures (DSMZ) and were handled with aseptic techniques according to general recommendations. Cells were kept in 25 - 300 cm^2 cell culture flasks with filter caps (TPP Techno Plastic Products AG) and incubated in a humidified incubator at 37°C and 5% (v/v) CO_2. MDA-MB231 cells were maintained in RPMI 1640 medium with stable glutamine (Merck Millipore, #FG1215) supplemented with 10% (v/v) heat inactivated fetal calf serum (hiFCS) (Biochrom, #S0615). KYSE150 cells were cultured in DMEM/Ham's F12 medium with stable glutamine (Merck Millipore, #FG4815) supplemented with 10% (v/v) hiFCS. HEK293, NIH3T3 and C2C12 cells were cultured in DMEM high glucose with stable glutamine, without sodium pyruvate (Thermo Fisher Scientific, #10566). For HEK293 and NIH3T3 medium was supplemented with 10% (v/v) hiFCS. For C2C12 medium was supplemented with 20% (v/v) hiFCS and 2.5 mg/L insulin (Sigma Aldrich, #I9278). FCS was heat inactivated for 30 min at 56°C in a water bath and sterile filtered. For continuous cell culture, cells were split 2-3 times a week to be subconfluent at the next round of splitting. C2C12 cells were split every second day. To split cells, old media was aspirated and cell layers were carefully washed once with sterile PBS (Merck Millipore, #L1825) and then covered with a thin layer of trypsin/EDTA solution (Thermo Fisher Scientific, #R001100) at RT. After 2-4 min cells were thoroughly resuspend in fresh medium and a fraction was transferred into a new flask. Cell numbers were determined with a Countess II Automated Cell Counter (Thermo Fisher Scientific, #AMQAX1000) according to the manufacturer's instructions. A549, A498, KYSE70, KPL-4, T47D, MCF7, EFM-19, ZR-75-1 cell lines for western blot analyses were kindly provided as snap-frozen cell pellets from colleagues and were handled according to ATCC recommendations.

3.2 Proliferation assay by Xcelligence

Proliferation of cell lines was measured using Xcelligence DP device (OMNI Life Science) according to the manufacturer's instructions. After seeding of 2000 cells/96-well in quadroduplicate, cells were permanently grown for 6 days. Eight hours after cell seeding, proliferation signals measured by conductivity were normalized to one for each individual well. Conductivity is represented as cell index, which is a compilation of cell division, cell growth, and cell adherence and served as surrogate for cell proliferation.

3.3 Differentiation of C2C12 cells

For differentiation of C2C12 myoblasts, cells grown to confluence and growth medium was changed to differentiation medium (DMEM with 2% (v/v) heat inactivated horse serum (HS) (Thermo Fisher Scientific, #26050088)). During the differentiation phase 90% of old medium was replaced by fresh medium every day.

3.4 Antibody generation

Customized methylation specific antibody SY46 was generated and purified by an external provider. For immunization of rabbit, a peptide corresponding to amino acids 361-380 from human p53 protein (NP_000537.3) with a mono-methyl lysine at the corresponding amino acid K370 (NH2-GSRAHSSHLK(me1)SKKGQSTSRH-COOH) was cross-linked at its N-terminus to Keyhole limpet hemocyanin (KLH) carrier protein. Antibodies were purified from sera of immunized rabbits via cross-affinity purification using a methylated and a non-methylated version of the peptide (see **Fig. 4.13**).

3.5 ELISA

Enzyme-linked immunosorbent assay (ELISA) for validation of SY46 methylation specificity was performed by the same provider as antibody generation using a sandwich ELISA. Briefly, 96-well plates were spotted with methylated or non-methylated peptide and blocked with BSA. Aliquots of different purification steps were incubated in a ten-fold dilution series with the peptides and unbound antibodies were washed off. Primary antibodies specific to the peptides were detected by secondary goat anti-rabbit antibodies cross-linked to horseradish peroxidase (HRP). Signals were measured at 492 nM after substrate addition.

3.6 IC_{50} determination

IC_{50} is the concentration of a compound at which a biological process is inhibited half maximal. For IC_{50} determination cells were treated with a half-logarithmic dilution series of the corresponding compound. For calculation of the IC_{50} measurement, values were normalized by setting the value at the highest compound concentration to 0% and the value at the lowest compound concentration to 100%. Normalized values were plotted against the log10-transformed compound concentrations and a regression curve was determined with GraphPad Prism 6 software using a four parameter non-linear dose-response model and least squares fit.

Equation: log(inhibitor) vs. response - Variable slope:

$$Y = Bottom + (Top - Bottom)/(1 + 10^{((LogIC_{50} - X) * HillSlope))}$$

X: log of dose or concentration
Y: response, decreasing as X increases
Top and Bottom: Plateaus in the same unit as Y
logIC_{50}: same log units as X
Hill slope: Slopefactor or Hill slope, unitless.

3.7 Proteom study

For detection of global lysine methylation in cells approximately $2x10^8$ MDA-MB231 cells stably overexpressing SMYD2-WT were harvested by 2-3 min detachment in Trypsin/EDTA solution and resuspension in cell culture media as described in **3.1**. Cells were centrifuged

for 5 min at 300x g and the cell pellet was washed twice in PBS. The dry cell pellet was snap-frozen and sent to an external provider. There, cell lysate preparation and enrichment of methylated peptides was done as described in[42]. For immuno-affinity enrichment either the customized SY46 methylation antibody or a pan-mono/di-methyl lysine antibody (Cell Signaling Technology, #14809) were used. Peptides were detected by LC-MS/MS. A list of all detected peptides containing methylated lysines was provided.

3.8 In silico analyzes

Copy number alterations (CNA) CNA for SMYD2 in human cancer patient samples were retrieved from cBioPortal database (`www.cbioportal.org/`).

Sequence motifs Sequence motifs were generated with WebLogo 3 software (`http://weblogo.threeplusone.com/create.cgi`) using the standard hydrophobicity color schema for amino acids.

Sequence alignments Sequence alignments were performed with Jalview (`http://www.jalview.org/`).

3.9 RNA microarray expression analyzes

For study of global gene expression of stable SMYD2 knockdown cells, cells were seeded into six-well culture plates in quadroduplicates and grown for 48 h until subconfluence. RNA was extracted using RNeasy Plus mini Kit (Qiagen, #74134) as described in **3.15**. For each replicate, 250 ng of total RNA was amplified using the GeneChip WT PLUS Reagent Kit (Affymetrix, #902281) according to the protocol described in User Manual Target Preparation for GeneChip Whole Transcript (WT) Expression Arrays (P/N 703174 Rev. 2). An Affymetrix Human Gene 2.1 ST 96-array plate was hybridized with 3 g of fragmented and labeled single stranded cDNA, washed, stained, and scanned according to the protocol described in the User Manual GeneTitan Instrument User Guide for Expression Arrays Plates (P/N 702933 Rev.1) and Affymetrix geneChip Command Console User's Guide (P/N 702569 Rev.9) using the Affymetrix geneTitan instrument. Principal component and correlation analyses were used to confirm data reproducibility. Differentially expressed probe sets were determined by carrying out paired t test comparisons of knockdown versus control cells. Significant probe sets with a FDR (Benjamini–Hochberg) < 0.1 were filtered by fold-change > 2.0 using Expressionist-GeneData software.

3.10 In vitro methylation

One µg of recombinant SMYD2 enzyme was incubated with 1 µg of protein substrate overnight at RT in 20 µL methylation buffer (50 mM Tris-HCl pH 9.0, 0.01% (w/v) BSA, 0.0022% (v/v) pluronic, 1 mM DTT, 100 µM SAM). Reactions were stopped by addition of 20 µL sample buffer and denaturation for 10 min at 95°C. Recombinant SMYD2 enzyme and recombinant p53 was kindly provided inhouse by Dr. Joerg Weiske and was generated as previously described[32]. Recombinant AHNAK-CRUs were expressed and purified as described in **3.33**.

3.11 NIH3T3 focus formation assay

For focus formation assays, 1x10^5 NIH3T3 cells were seeded into 6-well plates and transduced with retrovirus (for H-RAS-G12V) and/or lentivirus (for SMYD2-WT, SMYD2-H207A or GFP) as described in **3.32**. If needed, 24 h later cells were additionally treated with small molecule inhibitors. Media with/without fresh compound was replaced every second day. Focus formation was analyzed 10-20 days after transduction. For that, cells were washed one in PBS buffer and stained with methylene blue solution (Sigma-Aldrich, #77515) for 5 min. Cells were carefully washed twice with PBS and images were scanned for documentation. Retroviral H-RAS-G12V was kindly provided inhouse from Dr. Michael Steckel and was prepared from pRev(MSG)CMV-hRasG12V-Neo plasmid by the Natural and Medical Sciences Institute of the University of Tuebingen.

3.12 GATEWAY cloning

To efficiently transfer DNA sequences of interest into different plasmid vectors Gateway cloning was utilized[64]. Synthetic DNA fragments flanked by appropriate attB-sites were ordered from GeneArt (Thermo Fisher Scientific). The fragments were first cloned into a donor vector via BP reaction to generate an entry clone. Inserts were further transferred from an entry clone into the desired destination vectors via LR reaction to generate the corresponding expression clones. A list of vectors and their application is shown in **Tab. 3.2**. BP and LR reactions were performed according to manufacturer's instructions using BP clonase II (Thermo Fisher Scientific, #11789) and LR clonase II (Thermo Fisher Scientific, #117110) enzyme kits, respectively.

Tab. 3.2: List of GATEWAY compatible plasmids

Name	Company	Cat. #	Description
pDONR221	Thermo Fisher Scientific	12536017	donor vector for entry clones
pDEST15	Thermo Fisher Scientific	11802014	IPTG inducible expression vector for E. coli
pLenti6.3/V5Dest_verA	Thermo Fisher Scientific	V53306	mammalian expression vector for lentivirus production

3.13 Transformation of chemically competent E.coli

An aliquot of 50 µL of chemically competent E. coli was thawed on ice and up to 5 µL of sample or around 100 ng of plasmid DNA was added and incubated on ice for 30 min with occasional flicking. Cells were heat shocked for 45 s at 42°C in a thermo block, chilled on ice for 1 min and resuspended with 200 µL S.O.C. medium (Thermo Fisher Scientific, #15544034). Cells were shaken for 1 h, 180 rpm, 37°C. Depending on transformation efficiency, 20-100 µL was plated on LB agar containing the appropriate antibiotic. Colonies were picked after 24 h for further processing. For lenti-vectors the Stbl3 strain (Thermo Fisher Scientific, #C737303) was used. For protein expression the BL21 strain

(Thermo Fisher Scientific, #C600003) was used. Otherwise the DH10B derived Top10 strain (Thermo Fisher Scientific, #C404010) was used. LB medium was made of 1% (w/v) tryptone, 0.5% (w/v) yeast extract, 1% (w/v) NaCl, pH 7. LB agar was made of LB medium and 1.5% (w/v) bactoagar.

3.14 Plasmid DNA preparation

For small scale a single colony of transformed E. coli was inoculated in 4 mL LB-medium containing the appropriate selection marker (100 µg/mL ampicillin or 50 µg/mL kanamycin; Sigma Aldrich). Cultures were incubated overnight at 37°C, 180 rpm. Plasmid preparation was done with the QIAprep Spin miniprep Kit (Qiagen, #27104) according to the manufacturer's instructions. For transfection of mammalian cell cultures the EndoFree Plasmid Maxi Kit (Qiagen, #12362) was used according to the manufacturer's instructions.

3.15 RNA preparation

RNA from mammalian cell cultures was prepared with the RNeasy Plus mini Kit (Qiagen, #74134) according to the manufacturer's instructions. Typically, cells within a 6-well plate were lysed in 300 µL RLT Plus buffer. Sticky lysates were centrifuged for 1 min at 16,000x g through a QIAshredder spin column (Qiagen, #79656).

3.16 Preparation of cDNA

For preparation of cDNA from RNA the SuperScript III First-Strand Synthesis SuperMix (Thermo Fisher Scientific, #11752250) was used according to the manufacturer's instructions. Briefly, prior to cDNA synthesis RNA concentration from different samples was adjusted and 1-10 µg of total RNA was used for one 20 µL reaction. Samples were subsequently diluted with RNase-free water to 34 ng/µL of initial RNA and stored at -20°C until usage.

3.17 Quantitative real-time PCR with Taqmen probes

For RNA expression analyzes prepared cDNA was quantitatively measured by real-time PCR using a 7900HT Fast Real-Time PCR System (Thermo Fisher Scientific, #4329001). Ten µL reactions were prepared in MicroAmp Optical 384-Well Reaction Plates (Thermo Fisher Scientific, #4309849) using the TaqMan Fast Advanced Master Mix (Thermo Fisher Scientific, #4444557) according to the manufacturer's instructions and predesigned Taqmen probes (for individual probes see **Tab. 3.3**). Typically, 3 µL of pre-diluted cDNA (corresponding to 100 ng of input RNA) was used per reaction. Individual samples were analyzed in triplicates. Plates were sealed with MicroAmp Optical Adhesive Films (Thermo Fisher Scientific, #4311971). The thermal cycling profile was as follows: a single initial step of 2 min at 50°C and 20 s at 95°C, followed by 40 repeating cycles of 1 s at 95°C and 20 s at 60°C. Relative expression values were quantified by the ΔΔCt or Δ'Ct methods as described by Livak et al.[83].

Tab. 3.3: List of Taqmen probes from Thermo Fisher Scientific

Gene Symbol	Species	Assay ID	Reporter dye
SMYD2	Human	Hs00220210_m1	FAM
GAPDH	Human	Hs03929097_g1	FAM
ACTB	Human	Hs99999903_m1	FAM
Actb	Mouse	Mm00607939_s1	FAM
Ahnak	Mouse	Mm00613028_m1	FAM
Ahnak2	Mouse	Mm03015444_m1	FAM
Dmd	Mouse	Mm01216951_m1	FAM
Gapdh	Mouse	Mm99999915_g1	FAM
Mef2c	Mouse	Mm01340842_m1	FAM
Myod1	Mouse	Mm00440387_m1	FAM
Myog	Mouse	Mm00446194_m1	FAM
Rpl41	Mouse	Mm02524711_g1	FAM
Rps24	Mouse	Mm01623058_s1	FAM
Smyd1	Mouse	Mm00477663_m1	FAM
Smyd2	Mouse	Mm00660598_m1	FAM
Smyd3	Mouse	Mm00510201_m1	FAM
Smyd4	Mouse	Mm00662052_m1	FAM
Smyd5	Mouse	Mm00523415_m1	FAM

3.18 Photometric quantification of nucleic acids

For quantification of nucleic acids (DNA/RNA) the extinction at 260 nM was measured with a NanoDrop 2000 (Thermo Fisher Scientific, #ND-2000) according to the manufacturer's instructions.

3.19 Preparation of protein lysates from mammalian cell culture

Typically, cell culture media was aspirated and cells were washed once in chilled PBS buffer. Cells were then detached from culture plates by a cell scraper resuspended in PBS buffer and transferred to a reaction tube. Cells were collected by centrifugation for 5 min at 300x g and resuspended by pipetting in approximately 5 volumes of lysis buffer containing 1x Halt Protease and Phosphatase Inhibitor Cocktail (Thermo Fisher Scientific, #78440) and 1 U/µL Benzonase endonuclease (Merck Millipore, #101656). RIPA buffer (Thermo Fisher Scientific, #89900) was used for whole cell lysates and IP buffer (Thermo Fisher Scientific, #87787) was used for immuno-precipitation. Lysates were homogenized with a Bioruptor Sonication device (Diagnode, #UCD-200) for 5 min at high settings according to the manufacturer's instructions. Homogenized lysates were incubated on ice for 30 min and insoluble cell debris was spun down for 10 min at 16,000x g. The supernatant was used for further proceedings.

3.20 Preparation of tissue samples from mice

Mice (strain C57BL/6N, female, 7 weeks old) were euthanized according to regulatory requirements and different organs were surgically removed, shortly rinsed in PBS buffer,

snap frozen in liquid nitrogen and stored at -80°C until further processing.

3.21 Preparation of protein lysates from mice tissue

Tissue samples were crushed in liquid nitrogen with a mortar and pestle and resuspended in tissue lysis buffer (50 mM HEPES pH 7.4, 150 mM NaCl, 1 mM EDTA, 1 mM EGTA, 0.1% (w/v) SDS, 1% (w/v) Triton-X100, 10% (v/v) glycerol, 1 mM DTT, 2x Halt Protease and Phosphatase inhibitor cocktail, 2 U/µL Benzonase). Tissue lysates were homogenized first by sonication as described in 3.19 and second after 20 min of incubation on ice by centrifugation for 10 min at 16,000x g through a QIAshredder spin column (Qiagen, #79656). Insoluble debris formed a fragile pellet at the bottom of the QIAshredder tube and the supernatant with soluble protein was transferred to a new reaction tube.

3.22 Determination of protein concentrations

To determine the protein concentration of cell/tissue lysates, an aliquot of the sample of 1-10 µL was added to 200 µL BCA reagent (freshly prepared working solution from 50 volumes reagent A and 1 volume reagent B, Thermo Fisher Scientific, #23225) in 96-well plates (TPP, #92696) and incubated for 30 min at 37°C. A defined BSA dilution series (Thermo Fisher Scientific, #23208) was used to create a standard curve. After 30 min the extinction at 562 nM was measured using a Sunrise plate reader (Tecan, #16039400). Values were blanked against BCA reagent with lysis buffer instead of protein samples and absolute protein concentrations were calculated from the BSA standard curve.

3.23 Immuno-precipitation

Protein lysates were obtained as described in **3.19**. An aliquot of 500 µL with 2 mg/mL total protein was used per IP in a 1.5 mL reaction tube. For mice tissue, original protein lysates were diluted in IP buffer to 2 mg/mL total protein. Two µg of primary antibody were given to the lysates and incubated on an overhead rotator overnight at 4°C. Isotype IgGs were used as negative controls. For precipitation of the immuno-complexes 50 µL of equilibrated Protein G dynabeads (Thermo Fisher Scientific, #10003D) were added and incubated with rotation for 1-2 h at 4°C. Dynabeads were recovered with a magnetic rack (Thermo Fisher Scientific, #12321D) and washed 3 times for 10 min with 1 mL of IP lysis buffer. Immuno-precipitated proteins were recovered by incubation of dynabeads for 10 min at 95°C in 50 µL 2x sample buffer (prepared with 1/2 vol. 4x LDS sample buffer (Thermo Fisher Scientific, #NP0007), 1/5 vol. of 10x NuPAGE Reducing Agent (Thermo Fisher Scientific, #NP0004) in aqua bidest). Samples were directly applied to protein gel electrophoresis.

3.24 Denaturing protein gel electrophoresis

Sixty five µL of concentration adjusted protein lysates were denatured with 25 µL 4x NuPAGE LDS Sample Buffer (Thermo Fisher Scientific, #NP0007) and 10 µL 10x NuPAGE Sample Reducing Agent (Thermo Fisher Scientific, #NP0004) for 10 min at 95°C. Samples were chilled for 1 min on ice and equal volumes were loaded onto a pre-cast 4-12% Bis/Tris NuPAGE Protein Midigel (Thermo Fisher Scientific, #WG1402A). Empty lanes

were loaded with 1x sample buffer. The gel was run on a XCell4 SureLoc Midi-Cell Electrophoresis System (Thermo Fisher Scientific, #WR0100) with 1x NuPAGE MOPS SDS Running Buffer (Thermo Fisher Scientific, #NP0001) at 80-150 V according to the manufacturer's instructions. A Precision Plus Protein Dual Color Standard (Biorad, #1610374) was used as size marker.

3.25 Coomassie staining of protein gels

To visualize protein bands after gel electrophoresis gels were stained with Coomassie InstantBlue (Expedeon, #ISB1L) according to the manufacturer's instructions.

3.26 Western transfer and immunodetection

Western transfer After separation of proteins by gel electrophoresis, proteins were transferred to nitrocellulose membranes (Thermo Fisher Scientific, #LC2009) by western blotting using a Trans-Blot Electrophoretic Transfer Cell (Biorad, #1703853) according to the manufacturer's instructions. The transfer was performed in 1x NuPAGE Transfer Buffer (Thermo Fisher Scientific, #NP0006) w/o methanol for 2 h at 40 V or overnight at 15 V in an ice cooled water bath. After electrophoresis proteins transferred to the membrane were visualized by ATX Ponceau S (Sigma Aldrich, #09276-6X1EA-F) staining for 5 min to monitor transfer efficiency.

Immuno-detection Membranes were blocked for 1 h at RT in blocking buffer (1x PBS buffer (Thermo Fisher Scientific, #AM9625), 5% (w/v) non-fat dry milk (Sigma Aldrich, #M4709), 0.1% (v/v) Tween-20 (Sigma-Aldrich, #P2287)). After blocking, membranes were incubated overnight at 4°C with primary antibodies diluted in blocking buffer (for details on antibodies see **Tab. 3.4**). The primary antibody solution was removed and membranes were washed three times for at least 10 min in washing buffer (1x PBS, 0.1% (v/v) Tween-20). Membranes were then incubated for 2 h at RT with secondary antibodies diluted in blocking buffer. Blots were again washed three times for at 10 min in washing buffer before detection of the antibody signals were performed. Fluorescence labeled and HRP-conjugated secondary antibodies were detected with the appropriate channels using an Odyssey Fc imaging device (LI-COR) according to manufacturer's instructions. Fluorescent signals were quantified with the corresponding software package, if needed. HRP signals were obtained with the SuperSignal West Femto Maximum Sensitivity Substrate (Thermo Fisher Scientific, #34094).

3.27 In-cell-western (ICW)

Cells were seeded and treated within 100 µL cell culture media in 96-well cell culture plates (TPP, #92096) depending on the individual experiment. At the experimental endpoint, the medium was completely aspirated without detaching the cells and cells were fixed for 20 min in 50 µL ICW-fixation buffer (freshly prepared by diluting 37% formaldehyde solution (Sigma Aldrich, #33220) 1:10 in 1xPBS buffer). After two washes with 100 µL PBS, cells were incubated for 15 min in permeabilization buffer (0.25% (v/v) Triton X-100 (Sigma Aldrich, #93443) in PBS buffer). After one wash with PBS, cells were incubated for 1 h at RT in 100 µL ICW-blocking buffer (5% (w/v) nonfat dry milk (Sigma Aldrich,

Tab. 3.4: List of primary antibodies for WB

Target	**MW (kDa)**	**Type/ Host**	**Company**	**Cat.#**	**Dilution**
β-Actin	42	mMs	Abcam	ab8224	1:1000
β-Tubulin	50	pRb	Santa Cruz Biotech.	sc-9104	1:1000
AHNAK (C-terminus)	629	mMs	Santa Cruz Biotech.	sc-390743	1:500
AHNAK (N-terminus)	629	mMs	Antikörper-online	ABIN263928	1:1000
c-myc tag	-	mMs	Santa Cruz Biotech.	sc-40	1:2000
GAPDH	35	mMs	Santa Cruz Biotech.	sc-32233	1:2000
H3	17	mMs	Abcam	ab10799	1:2000
H3 K36me1	17	pRb	Abcam	ab9048	1:1000
H3 K36me2	17	mRb	Cell Signaling Tech.	2901	1:1000
H3 K36me3	17	pRb	Abcam	ab9050	1:1000
H3K4me1	17	pRb	Abcam	ab8895	1:1000
H3K4me3	17	pRb	Abcam	ab8580	1:1000
HSP90	90	mMs	BD Pharmingen	610418	1:5000
IgG-isotyp	-	pMs	Santa Cruz Biotech.	sc-2025	-
IgG-isotyp	-	pRb	Santa Cruz Biotech.	sc-2027	-
Kme1 (SY46)	-	pRb	-	customized	1:250
MyoD	34	pRb	Santa Cruz Biotech.	sc-304	1:1000
myogenin	25	mMs	Santa Cruz Biotech.	sc-12732	1:1000
myosin (MHC)	223	mMs	Developmental Studies Hybridoma Bank	MF 20-c	1:1000
p53	53	mMs	BD Pharmingen	554294	1:1000
p-ERK1/2	44, 42	mRb	Cell Signaling Tech.	4370	1:1000
SMYD1	56	mMs	Santa Cruz Biotech.	sc-514805	1:1000
SMYD2	50	pRb	Abcam	ab108217	1:1000
SMYD2	50	mMs	Santa Cruz Biotech.	sc-393827	1:500
SMYD3	49	pRb	Santa Cruz Biotech.	sc-67210	1:1000
SMYD5	47	pRb	Thermo Fisher Scientific	PA5-29153	1:1000

Tab. 3.5: List of secondary goat antibodies

Target	Fluorophore	Company	Cat.#	Dilution	Application
Mouse	Alexa Fluor 680	Thermo Fisher Sc.	A-21058	1:2000	WB, ICW
Rabbit	Alexa Fluor 680	Thermo Fisher Sc.	A-21109	1:2000	WB, ICW
Mouse	IRDye 800CW	LI-COR	926-32210	1:2000	WB, ICW
Rabbit	IRDye 800CW	LI-COR	926-32211	1:2000	WB, ICW
Rabbit	Alexa Fluor 488	Thermo Fisher Sc.	A-11008	1:1000	IF
Rabbit	Alexa Fluor 555	Thermo Fisher Sc.	A-21428	1:1000	IF
Rabbit	Alexa Fluor 647	Thermo Fisher Sc.	A-21245	1:1000	IF
Mouse	Alexa Fluor 488	Thermo Fisher Sc.	A-11001	1:1000	IF
Mouse	Alexa Fluor 555	Thermo Fisher Sc.	A-21422	1:1000	IF
Mouse	Alexa Fluor 647	Thermo Fisher Sc.	A-21235	1:1000	IF
Rabbit	HRP-conjugated	Santa Cruz Biotech.	sc-2004	1:1000	WB
Mouse	HRP-conjugated	Santa Cruz Biotech.	sc-2005	1:1000	WB

#M4709) in PBS buffer). After blocking, cells were exposed overnight at 4°C to 50 µL primary methylation antibody (SY46, 1:200 in ICW-blocking buffer). One row of cells on each plate was not exposed to primary antibody and served as background control. The wells were washed three times for 10 min with 100 µL PBS buffer before secondary IR800-conjugated antibody (LI-COR, #92632211) and DNA-intercalating dye DRAQ5 (Thermo Fisher Scientific, #62251) were incubated for 3 h at RT (both 1:1000 in ICW-blocking buffer). After five washes with PBS, the fluorescence in each well was measured on an Odyssey CLx scanner (LI-COR) at 800 nM (SY46 methylation signal) and 700 nM (DRAQ5 signal) according to the manufacturer's instructions. Fluorescence intensity was quantified and normalized to background (wells with no primary antibody) and DRAQ5 signals.

3.28 Confocal fluorescence microscopy

Cells were seeded and treated within 200 µL cell culture media in an 8-well Millicell EZ chamber slide cell (Merck Millipore, #PEZGS0816) depending on the individual experiment. At the experimental endpoint, the medium was completely aspirated without detaching the cells and cells were fixed at RT for 10 min in 100 µL fixation buffer (freshly prepared by diluting 37% formaldehyde solution (Sigma Aldrich, #33220) 1:10 in 1x PBS buffer). After two washes with PBS, cells were incubated at RT for 1 h in LSM-blocking buffer (5% (v/v) goat-serum (Thermo Fisher Scientific, #16210064) and 0.5% Triton X-100 (Sigma Aldrich, #93443) in Odyssey blocking buffer (LI-COR, #P/N 927-40000)). After blocking, cells were exposed overnight at 4°C to primary antibodies diluted in LSM-blocking buffer (for details on antibodies see **Tab. 3.6**). Isotype control IgGs were used as negative controls. The wells were washed three times for 10 min with LSM blocking buffer before Alexa Flour-conjugated secondary antibodies (for details see **Tab. 3.5**) were incubated for 2 h at RT in LSM blocking buffer. After one wash, nuclei were stained for 10 min with DNA-intercalating dyes either DAPI (Thermo Fisher Scientific, #62251) or Hoechst 33342 (Thermo Fisher Scientific, #62249) diluted at 1 µg/mL in PBS. If desired, filamentous actin was stained with phalloidin A555 (Thermo Fisher Scientific, #A34055). A stock solution of 200 U/mL in methanol was diluted 1:30 in PBS and cells were incubated for 30 min at RT. After three additional washes for 5 min with PBS the well chamber was removed from the glass slide and samples were sealed with self-hardening VECTASHIELD

mounting medium (VECTRO Laboratories, #H-1400) and a coverslip according to the manufacturer's instructions. Images were acquired using an Zeiss LSM700 confocal microscope and a 63X, 1.3 N.A. oil-immersion objective according to the manufacturer's instructions.

Tab. 3.6: List of primary antibodies for IF/ICW

Target	**Host**	**Company**	**Cat.#**	**Dilution**
AHNAK (N-terminus)	mMs	Antikörper-online	ABIN263928	1:200
c-myc tag	mMs	Santa Cruz Biotech.	sc-40	1:500
IgG-isotyp control mouse	pMs	Santa Cruz Biotech.	sc-2025	1:200
IgG-isotyp control rabbit	pRb	Santa Cruz Biotech.	sc-2027	1:200
Kme1 (SY46)	pRb	-	customized	1:200
SMYD2	mMs	Santa Cruz Biotech.	sc-393827	1:200

3.29 Transient gene knockdown by siRNA transfection

For transient gene knockdown in cell culture, cells were seeded one day before transfection to reach a confluency on the day of transfection of around 80%. Transfection of siRNAs was done with the Lipofectamine RNAiMAX transfection reagent (Thermo Fisher Scientific, #13778-150) according to the manufacturer's instructions. For the generation of the transfection mix serum-reduced Opti-MEM medium (Thermo Fisher Scientific, #31985062) was used. For a 6-well typically 75 pmol of individual siRNAs was transfected at a lipid/siRNA ratio of 2:1. Cells were analyzed 3 days after transfection. For details on individual siRNAs see **Tab. 3.7**.

Tab. 3.7: List of siRNAs

Target	**Species**	**Company**	**Cat.#**
AHNAK	Human	Sigma Aldrich	EHU058201
SMYD2	Human	Sigma Aldrich	EHU049591
Renilla Luciferase (neg. control)	Renilla	Sigma Aldrich	EHURLUC

3.30 Plasmid transfection in mammalian cells

Cell were seeded one day before transfection to reach a confluency on the day of transfection of around 80%. Transfection of plasmids was achived with the Lipofectamine LTX PLUS Reagent (Thermo Fisher Scientific, #15338-100) according to the manufacturer's instructions. For the generation of the transfection mix serum-reduced Opti-MEM medium (Thermo Fisher Scientific, #31985062) was used. For a 6-well typically 2 µg of plasmid DNA was transfected at a lipid:DNA ratio of 3:1. For details on individual plasmids see **Tab. 3.2**.

3.31 Lentivirus production

All work was performed in an S2 laboratory. For lentivirus production 1.5^10 HEK293FT cells (Thermo Fisher Scientific, #R70007) were seeded in a 75 cm^2 cell culture flask (Corning, #3290) in DMEM high glucose medium (PAA, #E15-843) containing 10% FCS. The day after, cells were transfected with 2.4 µg of lenti-vector and 7 µg of ViraPower Lentiviral Packaging plasmid mix (Thermo Fisher Scientific, #K497500) using Lipofecatmine 2000 reagent (Thermo Fisher Scientific, #11668019) according to the manufacturer's instructions. The medium was exchanged after 24 h and cells were incubated for further 24 h This time, medium with virus particles was collected in a 50 mL reaction tube and stored at 4°C. Cells were incubated with fresh media for additional 24 h. Medium from both days was pooled and centrifuged for 10 min at 500x g at 4°C to remove cells. The virus particles containing supernatant was sterile filtered (pore size 0.45 µm) and ultra-centrifuged at 50,000x g for 2.5 h The virus pellet was resuspended in 400 µL medium and aliquots of 100 µL were stored at -80°C until transduction. Virus titer was determined by HIV-1 P24 ELISA (Perkin Elmer, #NEK050) according to the manufacturer's instructions. A list of shRNA clones for lentivirus production is shown in **Tab. 3.8**.

Tab. 3.8: List of shRNA clones for lentivirus production

Name	Target gene	Developer	TRC clone ID
sh#1	human SMYD2	Broad Institute	TRCN0000276083
sh#2	human SMYD2	Broad Institute	TRCN0000276082
sh#3	human SMYD2	Broad Institute	TRCN0000276154
sh#4	human SMYD2	Broad Institute	TRCN0000276085
sh#5	human SMYD2	Broad Institute	TRCN0000276155
control	-	Broad Institute	SHC201

3.32 Lentivirus transduction

Cells that were seeded in their appropriate cell culture medium in 12-well plates 24 h prior transduction to reach a confluency of 80-90% at the day of transduction. For transduction, the media was discarded and 250 µL fresh media containing pre-diluted lenti-viral particles was applied to the cells. Normally, cells were transduced at a MOI of 10. For calculations of MOI it was assumed that 1 in 10 particles is infective. For instance, 10 M viral particles (determined by p24 ELISA) were used to transduce 100,000 cells. Cells were incubated with the viral media for 6 h. The medium was then discarded and cells were trypsinized and transferred to 25 cm^2 cell culture flasks and incubated in fresh media for 2-3 days. After this recovery period, positive transduced cells were selected for at least 10 days with appropriate selection media. During the selection medium was further supplemented with 1x penicillin-streptomycin (Thermo Fisher Scientific, #15140122) to prevent contamination.

3.33 Recombinant protein expression in E. coli of GST-tagged proteins and purification

In order to produce sufficient amounts of recombinant proteins, the targets of interest were cloned in an expression plasmid (for details on plasmids see **Tab. 3.2**) to generate an

N-terminal GST-fusion protein using GATEWAY cloning as described in **3.12**. Approximately 100 ng of expression plasmid was transformed into One Shot BL21(DE3) chemically competent E. coli (Thermo Fisher Scientific, #C600003) as described in **3.13**.

3.33.1 Small scale protein expression for optimization

Protein expression In order to identify well expressing clones and to determine optimized culture conditions for upscale protein expression, several colonies were picked the day after transformation, inoculated in a 2 mL LB-medium containing 200 µg/mL ampicillin (Sigma-Aldrich, A5354) in a 14 mL Falcon tube (Corning, prod. #352059) and incubated in a round shaker at 180 rpm, 37°C. After 5 h 50 µL of the pre-culture was transferred into 2 mL of freshly prepared 1x M9 medium in a 24-deep well plate (Qiagen, #19583), sealed with breathseal foil (Greiner Bio-One, #676051) and incubated over night at 180 rpm, 37°C. An aliquot of the pre-culture was prepared as glycerol stock by 1:1 dilution in 100% glycerol and stored at -80°C. The next day, protein expression was induced by adding 500 µL 5x circle growth medium (MP Biomedicals, #3000-146) incl. 200 µg/mL ampicillin and IPTG (Sigma-Aldrich, #11284) to obtain a final concentration of either 0, 0.1, 0.5, or 1.0 mM IPTG. This results in occupation of four wells/plate for each clone. Plates were prepared in duplicate and incubated for 20 h, 180 rpm and either 17°C or 27°C.

Cell disruption and sample preparation After 20 h of induction of protein expression, cells were harvested by centrifugation (Eppendorf, 5810R) within the 24-deep well plates for 15 min, 17°C, 4000 rpm. Supernatants were discarded and dry cell pellets were stored at -80°C until further processing. For cell lysis pellets were thawed on ice and resuspended in 500 µL lysis buffer (100 mM KPO, 300 mM NaCl, 10% (v/v) glycerol, 1x complete protease inhibitor cocktail (Roche, prod. #11697498001), 5 mM DTT, pH 7.5). Cell supspensions were then disrupted by sonication with a Covaris device (Covaris E210) according to the manufacturer's recommendations. After sonication lysates were diluted by additional 500 µL of lysis buffer and completely transferred to a 2 mL 96 deep well plate (Greiner Bio-One, ref. #786201). Cell debris was removed by centrifugation for 30 min, 4000 rpm, 4°C. Seven hundred fifty µL of supernatant was transferred to a 96-well turbo filter plate (Qiagen, #120025), which was put on a 96 deep well plate and again centrifuged for 2 min, 500x g, 4°C. The flow through was considered the soluble protein fraction.

GST-purification For GST-purification 100 µL of glutathione sepharose 4B (GE Healthcare, #17075601) was added to a 96-well receiver plate (Machery / Nagel, #89807) and centrifuged for 2 min, 500x g, 4°C. Flowthrough was discarded and beads were equilibrated by 500 µL washing buffer (identical to lysis buffer, w/o protease inh.) and centrifugation twice. Beads were then incubated with 600 µL soluble lysate, plates were sealed and shaken overhead for 30 min, 4°C. Plates were then centrifuged for 10 min, 100x g, 4°C. Beads were washed twice with 500 µL washing buffer and centrifuged 2 min, 500x g, 4°C. Bead bound proteins were recovered by overhead shaking for 30 min in 100 µL elution buffer (washing buffer containing 15 mM freshly prepared glutathione (Sigma-Aldrich, #G4251) at 4°C and centrifugation for 4 min, 100x g, 4°C. Final eluates were snap frozen in liquid nitrogen and stored at -80°C until further downstream analyses.

3.33.2 Midi scale protein expression for downstream analyses

protein expression For midi-scale protein expression the reagents and bacteria cultures were scaled up from small-scale proportionally to reach a final culture volume of 400 mL. Glycerol stocks of positive screened clones were used as starting material. Optimized IPTG concentration and temperature were chosen based on small scale results.

Cell disruption and GST-purification Cell pellets were resuspended in 5 mL lysis buffer and transferred to TC20 ultra centrifuge tubes (Biosciences, #520012). Cells were disrupted by sonication as before and lysate volume was adjusted to 10 mL with lysis buffer and ultracentrifuged for 45 min, 25,000x g, 4°C. The supernatant applied to 500 µL equilibrated glutathione sepharose 4B beads.

GST-purification GST-purification was basically done as in small-scale. Five hundred µL of glutathione sepharose 4B transferred to a poly-prep chromatography column (Biorad, #7311550) was used per 400 mL cell culture. Equilibration was done with 2x 6 mL wash buffer and centrifugation for 3 min, 150x g, 4°C. For protein binding, the supernatant after ultracentrifugation was incubated with equilibrated beads by overhead shaking of 30 min at 4°C. Flowthrough was discarded and beads were washed twice with 6 mL washing buffer. GST-fusion proteins were eluted by overhead shaking for 30 min in 750 µL elution buffer for 30 min at 4°C following centrifugation for 5 min, 300x g, 4°C. Elution step was repeated three times and eluates were pooled subsequently.

Removal of free glutathione Free glutathione in eluates was removed by size exclusion chromatography using a zeba spin desalting column (Thermo Fisher Scientific, #89892). The column was equilibrated with 2 mL washing buffer and centrifugation for 2 min, 1000x g, 4°C. The step was repeated four times before 2 mL of the eluate was applied and centrifuged as before. The final flow through was recovered and aliquots were snap frozen in liquid nitrogen and stored at -80°C.

3.34 Biochemical methylation assay for western blot analyses

In vitro methylation of recombinant proteins for western blot analyses was performed over night at RT in methylation buffer (50 mM Tris/HCl pH 9.0, 1 mM DTT, 0.01% (w/v) BSA, and 0.0022% (v/v) Pluronic) plus 1 µM freshly prepared S-adenosyl-L-methionine (SAM) and 1 µg of recombinant SMYD2 enzyme as well as 1 µg of recombinant substrate protein. Reactions were stopped by addition of sample buffer, denatured for 10 min at 95°C, subjected to SDS-PAGE and analyzed by Coomassie stain and western blot. Protein methylation was detected by western blot using SY46 methylation specific antibody.

3.35 Biochemical SPA methylation assay using oligo-peptide substrates and 3H-SAM

In vitro methyltransferase activity of SMYD2 towards oligopeptides was analyzed using a scintillation proximity assay (SPA), which measured methylation by the enzyme of synthetic, biotinylated peptides. For p53 a peptide derived from the C-terminal regulatory

domain comprising the K370 target lysine (K*) was derived (Biotin-Ahx-GSRAHSSHLK*SKKGQSTSRH-amide (Biosyntan)). For AHNAK, a peptide was derived from CRU domains comprising the suspected LK*GPK* target motif (PDVDLHLKGPKVKGD-Ahx-Biotin). A recombinant human SMYD2 full-length enzyme was kindly provided by Dr. Joerg Weiske and Dr. Amaury Ernesto Fernandez-Montalvan[32]. Assays were conducted in 384-well microtiter plates in methylation buffer (50 mM Tris/HCl pH 9.0, 1 mM DTT, 0.01% (w/v) BSA, and 0.0022% (v/v) Pluronic), and a final volume of 5 µL. The SMYD2 concentration in the assay was 3 nM, while tritiated S-adenosyl-L-methionine (3H-SAM, Perkin Elmer, #NET155H) and the peptide substrate were present at 60 nM and 1 µM, respectively (apparent Km of both substrates). Enzyme kinetics were followed over 2 h by quenching the reactions as described above at time points 0, 5, 10, 15, 30, 60, 90, and 120 min. Compounds were tested in 11-point, 3.5-fold dilution series ranging from 0.1 nM to 20 µM. Reactions were run for 30 min and quenched by adding Streptavidin PS SPA imaging beads (PerkinElmer) to a concentration of 3.12 µg/L and 25 µM "cold" SAM (Sigma, #A-2408). The amount of product was evaluated using a Viewlux (PerkinElmer) CCD plate imaging device [emission filter 613/55].

4 Results

4.1 Establishment of a cellular mechanistic assay

4.1.1 SMYD2 shows frequent gene amplification

Apart from cancer indications with described relevance for SMYD2 like esophageal squamous cell carcinoma (ESCC), I was interested if SMYD2 showed mutations or copy number alterations (CNA) in other cancer indications. For that purpose, SMYD2 was queried at the cBioPortal web service for cancer genomics (`http://www.cbioportal.org/`) using data from 147 cancer studies. Using a cut-off of at least 2% of samples per study with SMYD2 alterations, 34 studies remained. Overall, the main alteration found for SMYD2 was gene amplification, which may indicate a potential oncogenic activity in contrast to loss of function alterations, which rather indicate tumor suppressor activity. Interestingly, 6 out of 9 breast cancer studies showed amplification of SMYD2 above the cut-off. Especially a large study of advanced breast cancer (METABRIC, Nature 2012[23]) detected SMYD2 amplification in nearly 25% of the samples, indicating potential significance of SMYD2 in that indication (**Fig. 4.7**).

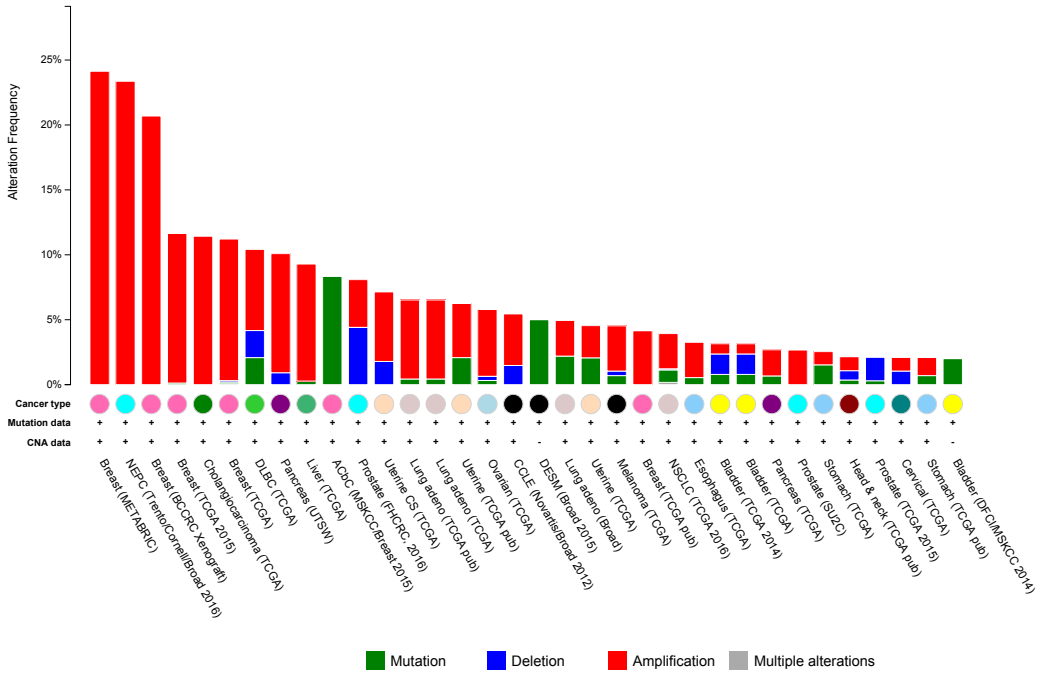


Fig. 4.7: Analysis of SMYD2 alterations in cancer using data from 147 cancer studies available at cBioPortal (http://www.cbioportal.org). Out of 147 studies available comprising different cancer indications 34 studies revealed SMYD2 gene alterations in at least 2% of the samples. From all alterations, gene amplification of SMYD2 was dominant, suggesting a potential oncogenic activity. Amplification of SMYD2 was observed in several indications, especially in invasive breast carcinoma (6 out of 9 breast cancer studies available >2% of samples).

4.1.2 SMYD2 is expressed in different cancer cell lines

For the development of a mechanistic assay that is able to measure SMYD2 activity, cancer cell lines with reasonable SMYD2 activity had to be identified. In the case of PKMTs in general, so far the best predictor for methylation activity is protein abundance. Therefore, several cell lines available inhouse were tested for SMYD2 expression by western blot, with a focus on breast cancer cell lines due to above received results from cBioPortal (**Fig. 4.7**). Specifically, analyzed cancer cell lines were derived from lung (A498, A549), cervix (HeLa), esophagus (KYSE70, KYSE150), and breast (KPL-4, MDA-MB231, T47D, MCF-7, EFM-19, ZR-75-1). SMYD2 protein was readily detected in HeLa and KYSE150 cells and in all breast cancer cell lines tested **Fig. 4.8**. High expression of SMYD2 in KYSE150 was in line with reported gene amplification in that cell line[70]. High expression of SMYD2 in all breast cancer cell lines in addition to frequent gene amplification in samples of breast cancer patients further pointed towards biological relevance of SMYD2 in that indication. For further analyses I selected KYSE150, MDA-MB231, and HeLa cell lines, because of their high expression of SMYD2 and their convenience in cell culture. Furthermore, the MDA-MB231 breast cancer cell line was reported to be sensitive to SMYD2 knockdown[91].

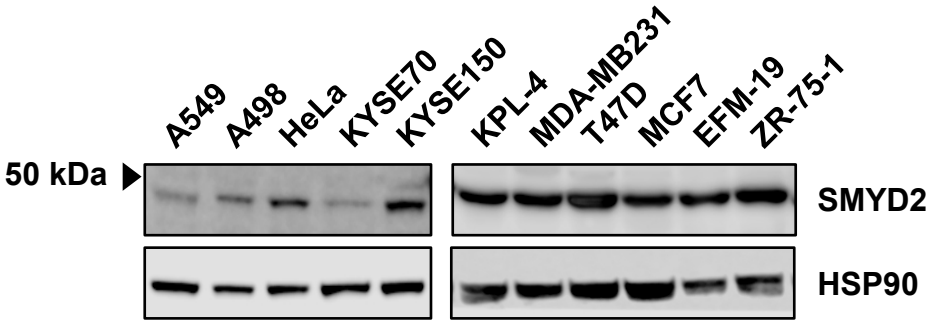


Fig. 4.8: SMYD2 is expressed in different cancer cell lines. Based on previous reports and data from cBioPortal several cancer cell lines derived from lung (A498, A549), cervix (HeLa), esophagus (KYSE70, KYSE150), and breast (KPL-4, MDA-MB231, T47D, MCF-7, EFM-19, ZR-75-1) were tested for SMYD2 protein expression by western blot. For SDS-PAGE, 30 µg total protein was loaded. Snap frozen cell pellets from all tested cell lines were kindly provided by colleagues.

4.1.3 SMYD2 does not influence global levels of histone methylation

SMYD2 was initially described to methylate histone H3 at lysine K36[13] and at lysine K4[1]. Therefore, I wondered, if a knockdown of SMYD2 will lead to changes in global histone methylation at lysine K4 and K36. For this purpose, MDA-MB231 breast cancer cells, which showed strong expression of SMYD2 protein in previous experiments, were transfected with a SMYD2 specific siRNA and histone methylation of H3K4me1, H3K4me3, H3K36me1, H3K36me2 and H3K36me3 were tested by western blot (**Fig. 4.9**). However, despite significant down-regulation of SMYD2, global methylation of tested histone marks was not affected. Therefore, I concluded that histone methylation of H3 K4 and K36 was not suitable for further assay development, at least under the tested conditions.

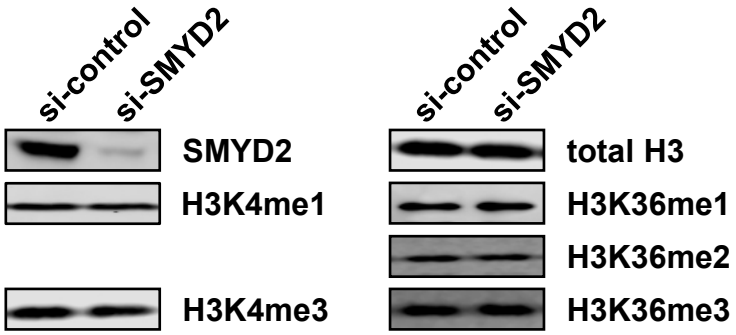


Fig. 4.9: Histone methylation is unaffected by SMYD2 knockdown.
MDA-MB231 cells were transfected with control siRNA or SMYD2-specific siRNA and proteins lysates were analyzed 3 days after transfection to evaluate SMYD2 knockdown and global histone methylation.

4.1.4 SMYD2 depletion does not change global gene transcription

Several reports have indicated that SMYD2 is involved in the regulation of transcription activity[1;19;27;105]. Identifying target genes whose transcriptional activity is regulated by SMYD2 would allow to establish a mechanistic assay. Therefore, I established stable SMYD2 knockdown cells from the MDA-MB231 breast cancer cell line model. Five different polyclonal cell lines were generated by lentiviral transduction of five independent shRNAs targeting different regions within the open reading frame of SMYD2. After lentivirus transduction, cells were kept in selection medium for ten days to generate stable polyclonal knockdown cell lines. Quantitative real time PCR data confirmed efficient knockdown of SMYD2 on transcript level of all five cell lines ranging from 70% (sh#4) to 95% (sh#2) knockdown efficiency compared to empty vector control cells (**Fig. 4.10A**). This translated also in strongly reduced protein levels of SMYD2 in the knockdown cell lines (**Fig. 4.10B**). However, in contrast to a previous report[91], proliferation of individual established knockdown cell lines was not significantly affected compared to the empty vector control cell line (**Fig. 4.11**).

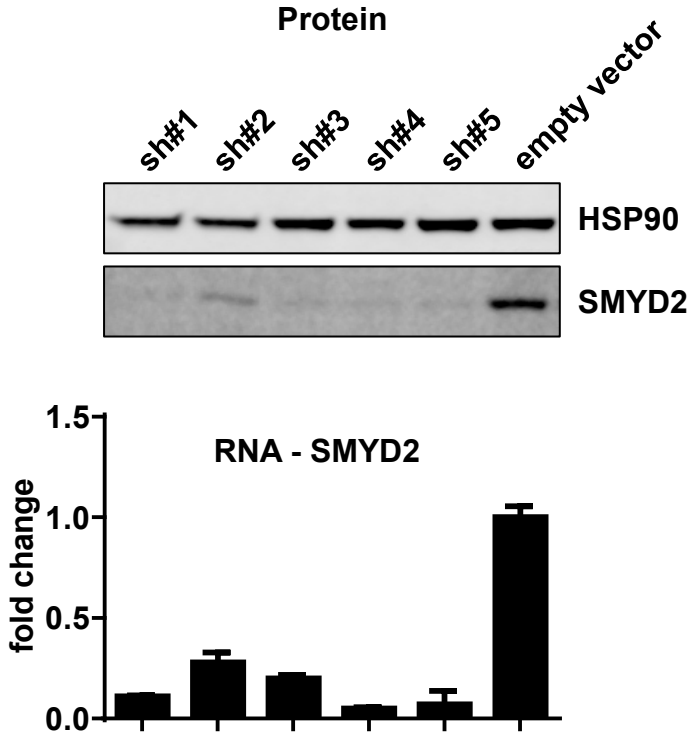


Fig. 4.10: Transduction of lentivirus shRNAs effectively reduced SMYD2 RNA and protein level.
MDA-MB231 cells were transduced with different lentiviral shRNA vectors against SMYD2 or empty vector control virus and tested for SMYD2 knockdown by qRT-PCR (A) and western blot (B) 10 days after puromycin selection.

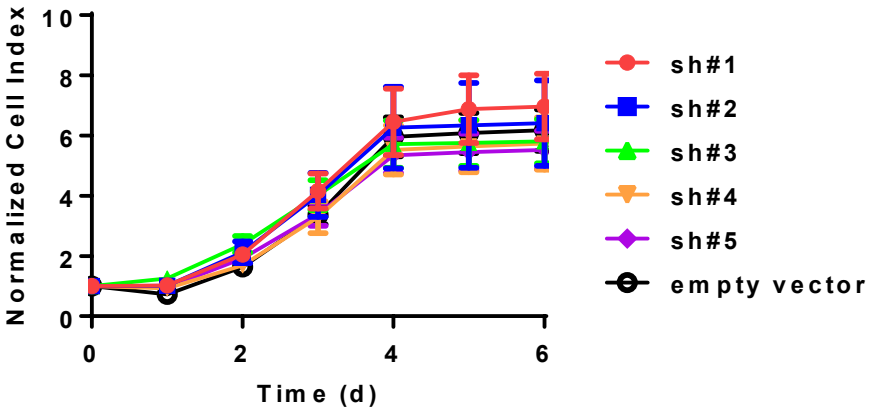


Fig. 4.11: Proliferation of MDA-MB231 cells after knockdown of SMYD2 is not affected. Proliferation was measured in MDA-MB231 cells stably transduced with SMYD2 specific shRNAs or empty vector control using Xcelligence device. After seeding 2000 cells/96-well in quadroduplicate, cells were permanently grown for 6 days. Eight hours after cell seeding proliferation signals were normalized to one for each individual well. Quantified data are represented as cell index, which is a combined measure of cell division, cell growth, and cell adhesion to the culture dish and served as surrogate for cell proliferation.

The four most efficient knockdown cell lines (sh#1, 3, 4, 5) as well as the empty vector control cell line were used to measure their global gene expression in an expression profiling experiment. Statistical testing with a principal component analysis indicated that individual replicates of the different cell lines clustered together, which confirmed high consistency of the prepared samples. However, cell lines derived from individual shRNA constructs did not separate in a consistent way from the control cell line indicating that the overall changes in gene expression were dominated by off-target effects independent from SMYD2 knockdown (**Fig. 4.12A**). As all shRNA constructs showed efficient knockdown of SMYD2, potential target genes regulated by SMYD2 should be affected in all knockdown cell lines. Therefore, gene expression in the different knockdown cell lines were compared to empty vector control cells. Differentially expressed genes were defined as having at least a two-fold change in gene expression with a false discovery rate of 10% (q BH <0.1). Using these criteria, the cell line derived from shRNA1 (sh#1) revealed nine differentially expressed genes. In contrast, the cell line derived from shRNA3 (sh#3) showed 152 differentially expressed genes, whereas cell lines derived from shRNA4 and 5 (sh#4 and sh#5) revealed 18 and 56 differentially expressed genes, respectively. When comparing the overlap of the different gene sets, SMYD2 was the only gene that was commonly regulated by all shRNAs (**Fig. 4.12B**). This indicated that the observed effects of the individual shRNAs on gene expression are most likely not connected to the SMYD2 knockdown and are rather off-target effects of the individual shRNAs used. Based on these data I concluded that SMYD2 is not a regulator of gene transcription in MDA-MB231 cells at least under the tested conditions and the chosen significance criteria and that mechanistic assay development based on changes in gene expression downstream of SMYD2 activity in not feasible.

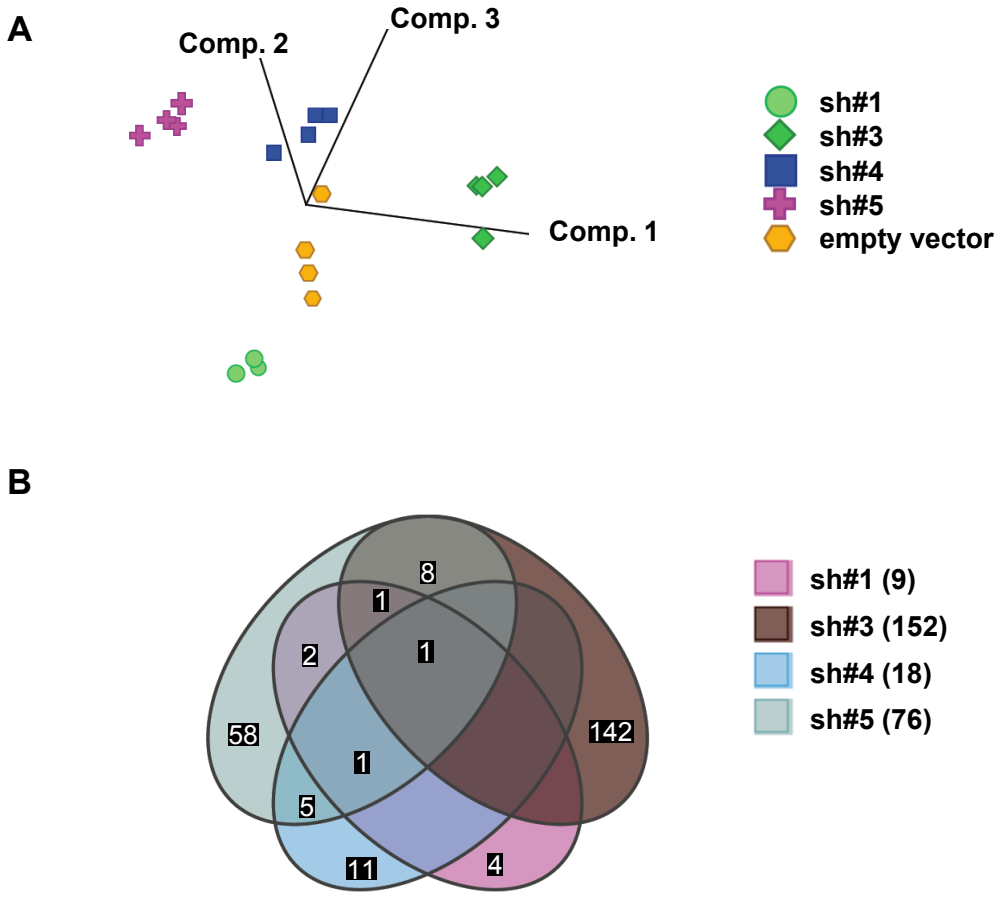


Fig. 4.12: SMYD2 knockdown has no significant impact on transcriptional regulation
Stable MDA-MB231 SMYD2 knockdown cells were established by lentiviral transduction of five independent shRNA constructs as well as an empty vector control cell line. RNA samples of four replicates were prepared for a human Affymetrix GeneChip array. Data were analyzed using GeneData Expressionist software. Principal component analysis was used to visualize overall differences in gene expression (A). Venn diagram compares differentially expressed genes relative to empty vector between of all five shRNA cell lines (B). Differentially expressed genes were defined as $\geq$ 2-fold change in gene expression (FDR q-value (BH) $<$ 0.1).

4.1.5 Generation of a p53 specific methylation antibody

In addition to histone methylation and regulation of gene expression, the tumor suppressor p53 had been described as a target for mono-methylation by SMYD2 at lysine K370[55]. To follow up on this report, rabbit polyclonal antibodies were raised against p53K370me1 and cross-purified by an external provider (**Fig. 4.13**, for details see **3.4**). Aliquots from different steps of the purification process were further tested by the external provider for methylation specificity by ELISA in a ten-fold dilution series (**Fig. 4.14**). As expected, eluate 1 showed an increase in reactivity towards the methylated p53 peptide compared to serum, whereas flowthrough 1 (FT1) revealed no reactivity (**Fig. 4.14A**). Eluate 1 was further purified against non-methylated peptides to reveal FT2 that was specific only for methylated p53 (**Fig. 4.13 column 2**). This was demonstrated by ELISA, where FT2, in contrast to eluate 2, detected methylated p53 peptide but not the non-methylated form

(**Fig. 4.14B,C**). For subsequent experiments antibodies derived from FT2 were used and are referred to as SY46 antibody (methylation specific antibody).

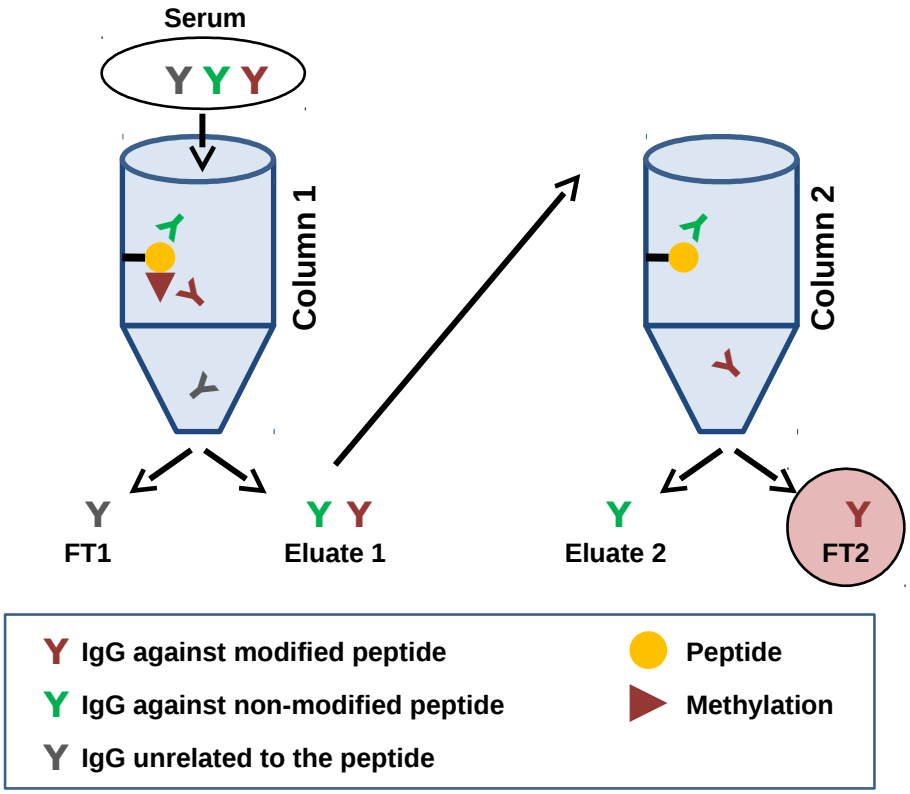


Fig. 4.13: Schematic representation of purification of a p53-methylation specific antibody
In order to detect p53 methylation, a polyclonal rabbit antibody was raised against a methylated p53 peptide. Serum from an immunized rabbit was purified via cross-affinity purification. Antibody generation and purification was done by an external provider.

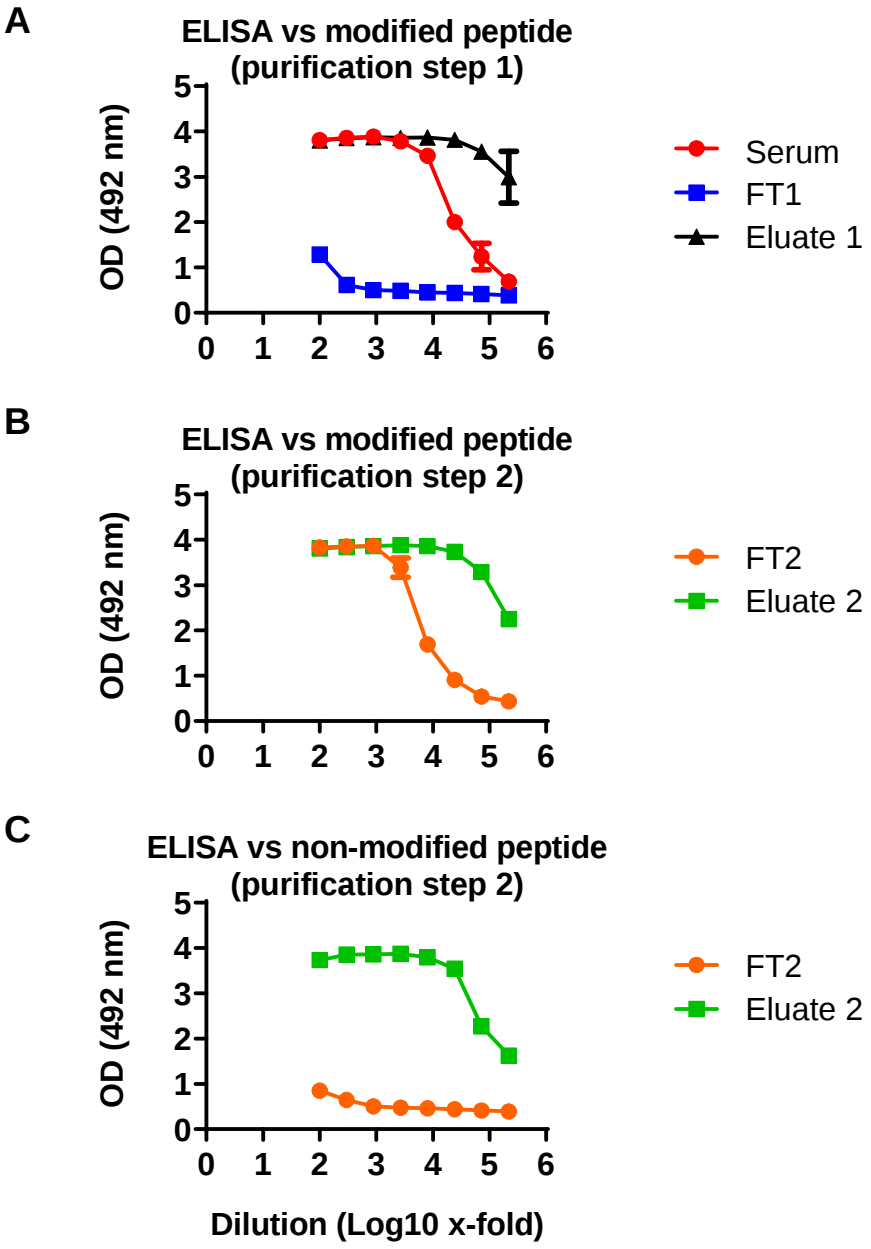


Fig. 4.14: ELISA analyses of different fractions of antibody purification confirms methylation specificity
Antibody fractions after cross-affinity purification were tested on ELISA for specificity towards methylated p53 peptide. Input serum was compared against flowthrough 1 (FT1) and eluate 1 (A). FT2 was compared with eluate 2 (B). FT2 and eluate 2 were tested for reactivity towards non-methyl-p53 peptid (C).

4.1.6 SY46 detects SMYD2 dependent methylation

To test, if SY46 can also detect methylation on full-length p53, recombinant p53 was incubated in a biochemical methylation reaction either in the presence or absence of recombinant SMYD2 and analyzed by western blot using SY46 antibody. Only p53 protein

incubated with SMYD2 was detected by SY46 (**Fig. 4.15**), suggesting methylation specific detection of p53 by SY46.

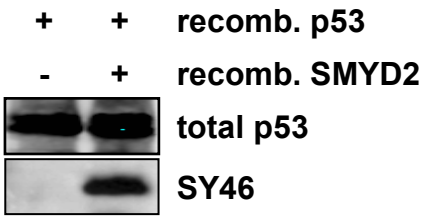


Fig. 4.15: Methylation antibody SY46 detects SMYD2-dependent methylation on full-length p53.
Recombinant full-length p53 was in vitro methylated by SMYD2 overnight and methylation was detected on western blot by SY46 methylation antibody.

To evaluate if p53 methylation also occurs in cells and if this can be detected with SY46 antibody I used the KYSE150 cell line model that has been previously described to overexpress SMYD2 due to gene amplification and also shows stabilized levels of p53[70]. Cells were treated with AZ505, a described SMYD2 tool compound[36], to evaluate, if p53 methylation can also be modulated by SMYD2 inhibition. Indeed, on western blot SY46 detected a protein band at the MW of p53 and the signal was reduced upon increasing concentrations of AZ505 (**Fig. 4.16**). This indicated that SMYD2 methylates p53 in cells. However, detection of endogenous p53 methylation was not strong enough for a quantitative medium throughput assay.

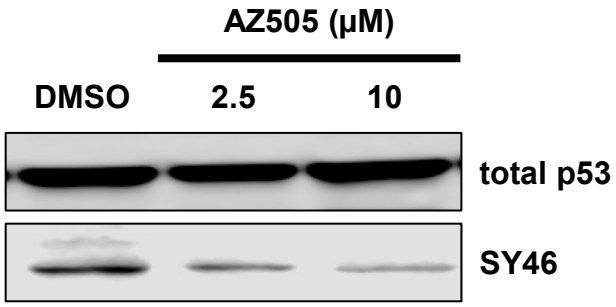


Fig. 4.16: SMYD2-dependent p53 methylation can be detected in cells.
KYSE-150 cells were treated with 2.5 or 10 µM of SMYD2 tool compound inhibitor AZ505 for 48 h and endogenous p53 methylation signal was analyzed on western blot with SY46 antibody.

Therefore, I intended to ectopically overexpress SMYD2 in order to increase the pool of methylated p53 relative to the total pool of p53. Surprisingly, ectopic overexpression of SMYD2 did not increase measurable p53 methylation, but revealed strong methylation signals (detected by SY46) on western above the size of the 100 kDa protein marker (**Fig. 4.17 asterisk**). The signal appeared to be either a result of methylation of several substrates or arising from degradation products of one large protein. The antibody SY46 also

reacted with a protein band at the expected size of histone H3 but this signal was not increased upon SMYD2 overexpression, indicating unspecific cross-reactivity of the antibody towards histone methylation (**Fig. 4.17 arrow**).

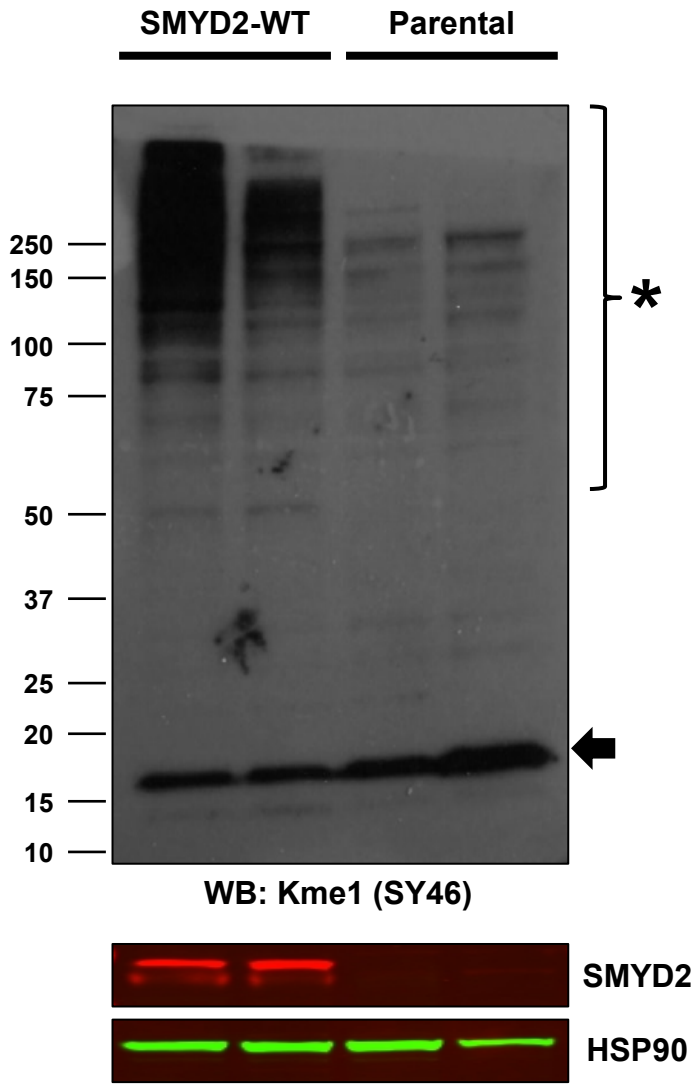


Fig. 4.17: Ectopic SMYD2 overexpression leads to massive increase in substrate methylation. KYSE150 cells were transfected with an SMYD2 overexpression plasmid and protein lysates were analyzed 48 h after transfection by western blot. SY46 antibody was used to detect methylated substrates. Numbers on the left indicate molecular weight of protein markers.

To check, if the strong signal detected by SY46 on western blot upon SMYD2 overexpression is specific and can be modulated by SMYD2 inhibition, KYSE150 cells transfected with a SMYD2 expression plasmid were treated with 10 µM AZ505 for up to 3 days and examined by western blot. The SY46 derived signal on western blot was decreased in samples of longer treatment (**Fig. 4.18 asterisk**). This indicated that SY46 in fact detects SMYD2-dependent protein methylation. In contrast, cross-reactivity towards a histone band was not affected by treatment (**Fig. 4.18 arrow**). In conclusion, SY46 detects protein methylation that is strongly induced upon ectopic SMYD2 overexpression and depends

on SMYD2 enzymatic activity.

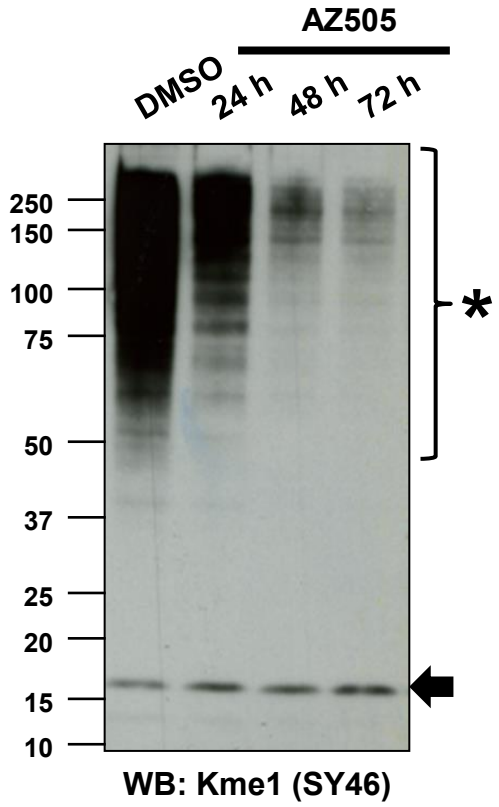


Fig. 4.18: Substrate methylation upon ectopic SMYD2 overexpression can be prevented by SMYD2 catalytic inhibition.
KYSE150 cells were transfected with SMYD2 overexpression plasmid and treated with 10 µM AZ505 for up to three days. Methylation was analyzed on western blot by SY46 antibody. Numbers on the left indicate molecular weight of protein markers.

To have additional assay options, I was wondering, if the strong methylation signal can also be detected in fixed cells via immuno-fluorescence staining. Therefore, KYSE150 cells were transfected with a SMYD2 expression plasmid, fixed 48 h after transfection and stained for ectopic SMYD2 expression (c-myc tag) and substrate methylation (SY46 antibody). As transfection efficiency was not 100%, SMYD2 positive and negative cells could be directly compared. Interestingly, ectopic SMYD2 was mainly localized to the cytoplasm, which probably could explain unaffected histone methylation in cells. Furthermore, SMYD2 positive cells also showed a strong cytoplasmic staining derived from SY46 antibody (**Fig. 4.19 arrowheads**), which tended to become enriched at the cell periphery in more confluent cells (**Fig. 4.19 center**). SY46 also stained nuclei in all cells, which is in line with the cross-reactivity of the antibody towards methylated histones. Interestingly, ectopic SMYD2 and its methylation substrate(s) co-localized to the spindle fibers in mitotic cells (**Fig. 4.19 yellow arrow**). This may point towards a role of SMYD2 during mitosis as a cytoskeletal component. Taken together, signals from SMYD2 dependent protein methylation could be detected by SY46 antibody by western blot and via immuno-fluorescence.

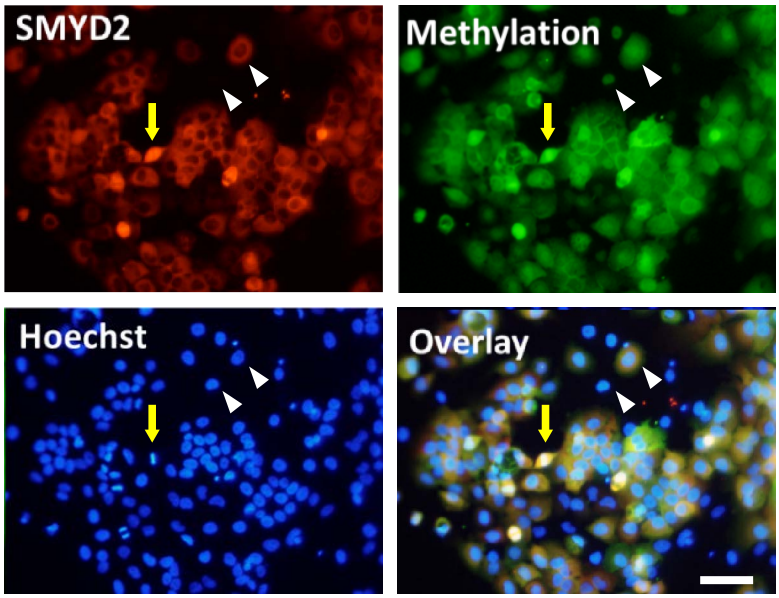


Fig. 4.19: SMYD2 substrate methylation occurs in the cytoplasm.
KYSE150 cells were transfected with SMYD2 overexpressing plasmids and prepared for immuno-fluorescence microscopy 48 h after transfection. Cells were stained for substrate methylation (SY46), SMYD2 via a c-myc tag, and Hoechst for nuclear DNA. Scale bar indicates 50 µM.

The strong methylation signals in immuno-fluorescence analyses suggested the possibility to measure the SMYD2 dependent methylation signal in a more quantitative in-cell-western. This assay is comparable to immuno-fluorescence staining but at a much lower resolution. In contrast to single cells, the overall fluorescence signal of a complete 96-well cavity is measured and wells with different treatment conditions can be quantitatively compared. Indeed, wells containing KYSE150 cells ectopically overexpressing SMYD2 revealed a stronger SY46 derived fluorescence signal compared to wells of parental cells. More importantly, inhibition of SMYD2 activity by increasing concentrations of inhibitor led to a dose-dependent reduction of detected methylation signals (**Fig. 4.20A**). For quantification purposes and as a control, the methylation signal derived from SY46 was normalized to DRAQ5 DNA stain, which served as a surrogate for cell number. Quantification data illustrate an approximately three-fold higher methylation signal over background (most likely histone methylation plus unspecific binding), which defined the assay window (**Fig. 4.20B**). AZ505 was used as a benchmark inhibitor for SMYD2 activity and was compared against two novel SMYD2 inhibitor candidates derived from the aminopyrazonline hit of initial screening campaign. The two compounds represented enantiomers of each other (**Fig. 4.21**). As expected for a pocket binder, one enantiomer named compound 1 showed essentially no methylation inhibition activity (the very weak inhibition observed at high concentrations is probably due to slow conversion between enantiomeres or residual inpurities from the other enantiomer) whereas the other enantiomer named compound 2 showed strong activity, which was even more potent than AZ505 (**Fig. 4.20B**).
The data demonstrated that the in-cell-western, which was based on ectopic SMYD2 overexpressing KYSE150 cells and a methylation specific antibody (SY46), provided a robust and reliable cellular mechanistic assay to quantitatively measure SMYD2 methylation ac-

tivity. Subsequently, this assay was used to determine cellular IC50 for SMYD2 inhibition of many chemical compounds (see Eggert et al.[32]). Among those, a compound similar to compound 2, but with an additional chlorine atom at the phenol residue, revealed strong cellular potency with an IC50 of 50 nM. Additional pharmacokinetic and pharmacodynamic testing in mice confirmed in vivo suitability of that compound and was therefore selected as preferred SMYD2 inhibitor probe referred to as BAY-598 (**Fig. 4.21** and Eggert et al.[32]).

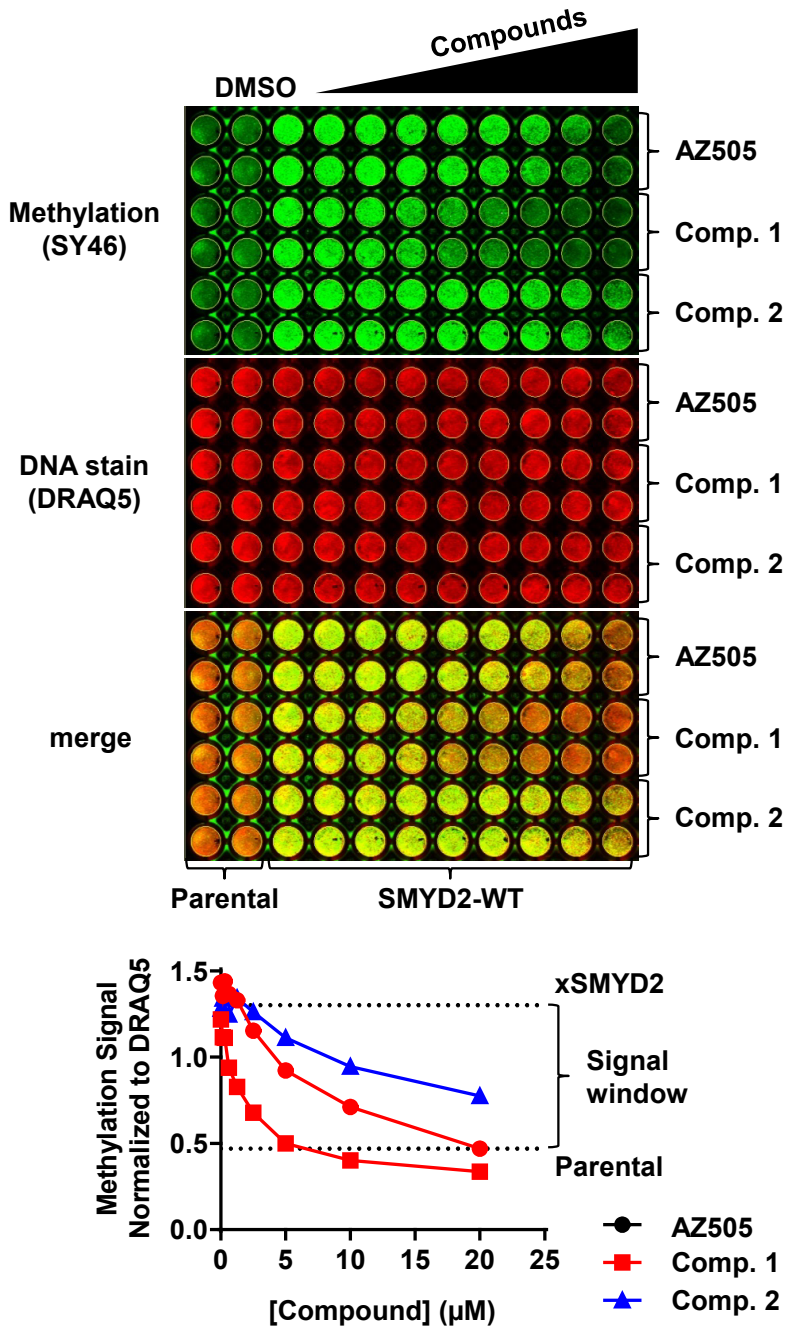


Fig. 4.20: SMYD2 substrate methylation can be quantified by In-Cell-Western detection. KYSE150 cells were transfected with SMYD2 overexpression plasmid and treated with a dilution series of novel SMYD2 inhibitor candidates for 72 h. AZ505 was used as benchmark. Substrate methylation was detected in an In-Cell-Western with SY46 antibody. DRAQ5 DNA staining was used for normalization. Methylation signals were quantified to compare different compounds.

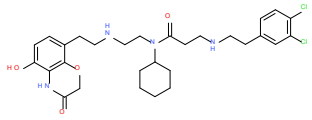
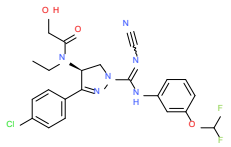
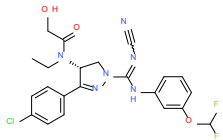
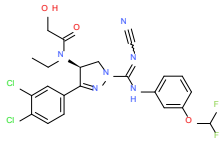
Structure	Name	SMYD2 biochem. IC50 (SPA)	SMYD2 cellular IC50 (ICW)
	AZ505	800 nM	3 µM
	Compound 1	400 nM	450 nM
	Compound 2	8 µM	18 µM
	BAY-598	30 nM	60 nM

Fig. 4.21: Summary of different SMYD2 small molecule inhibitors.

4.2 Discovery and validation of a novel SMYD2 substrate

Based on observations during the mechanistic assay development for SMYD2 it became clear that this enzyme possesses so far unexplored substrates. The identification of such substrates would probably allow for a deeper understanding of the biological function(s) of SMYD2.

4.2.1 SMYD2 dependent methylation is specific and occurs in several cell line models

To determine, if the observed methylation signal with SY46 antibody upon SMYD2 overexpression was a unique feature of the esophageal cancer cell line KYSE150 or a more general effect, additional cell lines derived from different tissues were tested. Therefore, I established several syngeneic polyclonal cell lines by stable lentivirus transduction of MDA-MB231 and HeLa. These cell lines ectopically overexpressed either wild type SMYD2 (SMYD2-WT) or SMYD2 mutated at amino acid residue 207 from histidine to alanine (SMYD2-H207A). This residue sits within the SAM binding domain and is supposed to render SMYD2 enzymatically inactive. In addition, cell lines were established, which overexpressed the close homolog SMYD3 either as wild type (SMYD3-WT) or as mutant form (SMYD3-N205A, similar to SMYD2 mutant). These two cell lines served as additional controls for the specificity of SMYD2 methyltransferase activity compared to other methyltransferase activities. Furthermore, GFP overexpressing cells served as additional control. Western blot analysis revealed successful overexpression of the corresponding proteins after two weeks of blasticidin selection (**Fig. 4.22**). The wild type proteins of the enzymes appeared to be stronger expressed as the corresponding mutants, probably, because the mutations were within the SAM binding pocket. This was supposed to prevent efficient SAM binding, which might have compromised protein stability. Importantly, similar to the KYSE150 cell line model, strong methylation signals were detected with the SY46 antibody at high molecular weights in the SMYD2-WT overexpressing clones of MDA-MB231 and HeLa (**Fig. 4.22**). No signals were detected in cell lines expressing the mutant SMYD2 protein, GFP, or the related SMYD3 enzymes. This demonstrated that the observed protein methylation is specific to SMYD2 enzyme activity and is not unique to the KYSE150 cell line and might have a functional role in different cell types.

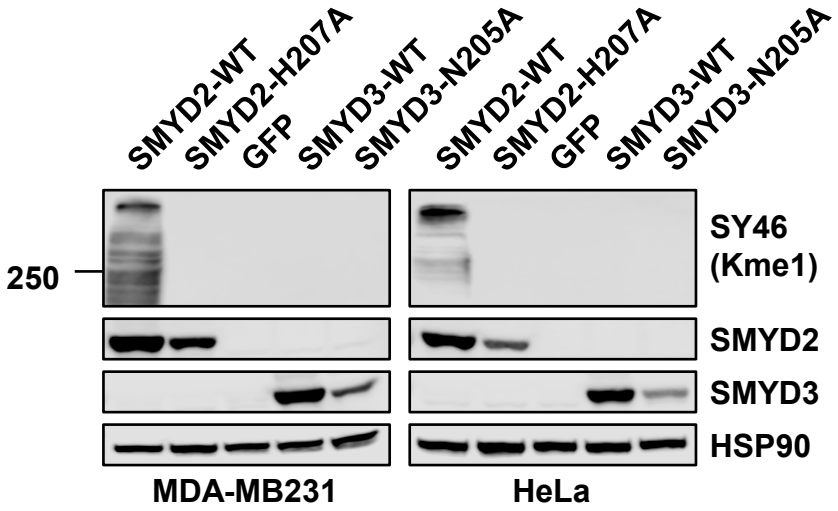


Fig. 4.22: SMYD2-dependent methylation appears in different cell lines.
Syngeneic MDA-MB231 and HeLa cell clones stably overexpressing the indicated proteins were analyzed by western blot. SY46 methylation is specific to SMYD2 and depends on SMYD2 enzymatic activity. An enzymatically dead point mutant H207A reveals no methylation signal. GFP and SMYD3 constructs served as additional control.

4.2.2 Proteomic approach detects many novel lysine methylation sites in SMYD2 overexpressing cells

The above results indicated that SY46 may be suitable to specifically capture methylated SMYD2 substrates. Previous studies have shown that it is possible to discover novel methylation sites on proteins from cell extracts in a more unbiased way by mass spectrometry[42]. Therefore, I aimed at following a similar proteomic approach. In principle, cells are lysed in denaturing urea buffer and proteins are proteolytically digested with trypsin into smaller peptides. Methylated peptides are then immuno-precipitated with methyl-specific antibodies and detected by LC-MS/MS. A schematic overview of the procedure is outlined in **Fig. 4.23**. For immunoprecipitation SY46 antibody and a commercial pan-Kme1/2 specific antibody were used. For the analysis the MDA-MB231 cell line stably overexpressing ectopic SMYD2-WT was chosen, because of the strong observed methylation signal(s). Snap-frozen cell pellets and SY46 antibody were sent to an external company, which performed sample preparations and MS analyses and a list of all detected methylated peptides was returned.

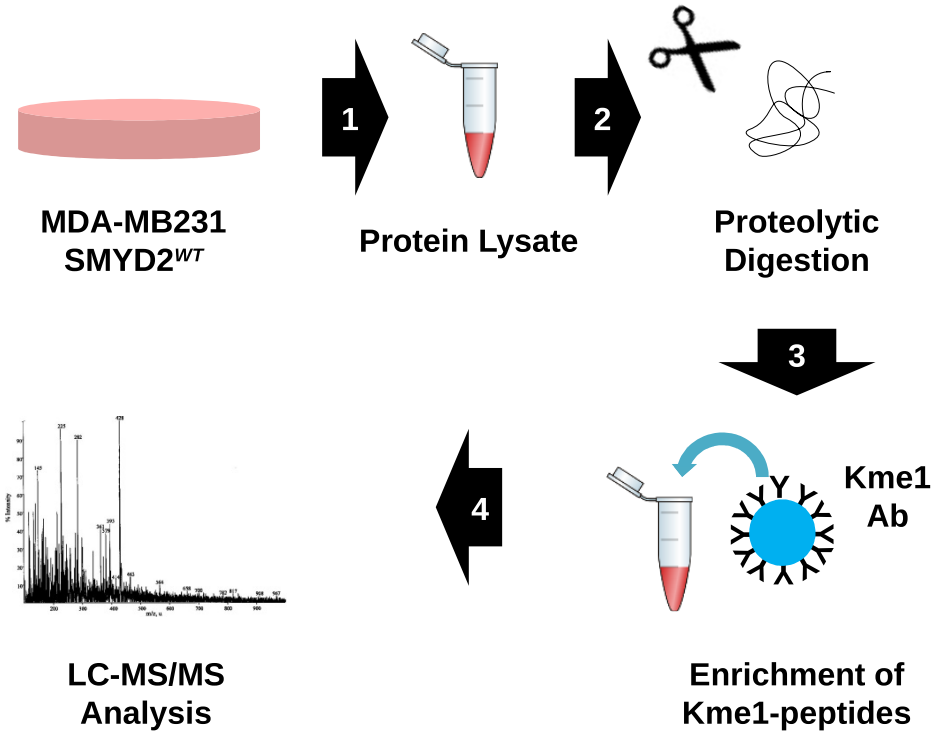


Fig. 4.23: Schematic representation of proteomic substrate identification approach. MDA-MB231 cells stably overexpressing SMYD2-WT were grown to obtain a cell pellet of approximately 2 mg. A snap-frozen cell pellet was then provided to an external company for downstream procedures (1-4). Cells were lysed in denaturing urea-buffer and digested by trysin to obtain shorter peptides. Lysine methylated peptides were enriched from whole cell extracts by immunoprecipitation using either SY46 antibody or a commercially available pan-Kme antibody. Enriched peptides were further detected by LC-MS/MS analysis. A list of all detected methylated peptides was provided.

Overall, capturing with SY46 antibody revealed 188 unique mono-methylation sites corresponding to 121 different proteins. Capturing with the pan-Kme1/2 antibody revealed 695 unique mono-methylation sites corresponding to 503 single proteins. The overlap of both antibodies was 118 methylation sites and 73 proteins, respectively (**Fig. 4.24A**). The difference in the number of detected methylation sites of both antibodies could be explained by different affinities to different methyl epitopes and by stronger constraints of the SY46 antibody towards certain methylation epitopes. Sequence logos of all detected methylation sites for both antibodies were generated and aligned to the p53 sequence around the K370 target lysine. Inspection of the sequence logos indicated an enrichment of leucine and phenylalanine at the -1 position. In addition, the fraction of methylated lysines detected by both antibodies also showed a clear enrichment for glycine and serine at the +1 position and for proline at the +2 position relative to the methylated lysine (**Fig. 4.24B**). The sequence motif derived from peptides captured by SY46 only partially matched the sequence of p53 peptide. This shows the ability of the antibody to capture a broader spectrum of methylated peptides. However, assuming a dominant impact of SMYD2 mediated protein methylation on the overall lysine methylation in the cell, it may also indicate a much broader spectrum of target motifs for SMYD2. However, the latter one is only a hypothetical assumption based on results from western blot and therefore, no clear conclusion could be made which of the peptides were methylated by SMYD2 and

which were methylated by other enzymes.

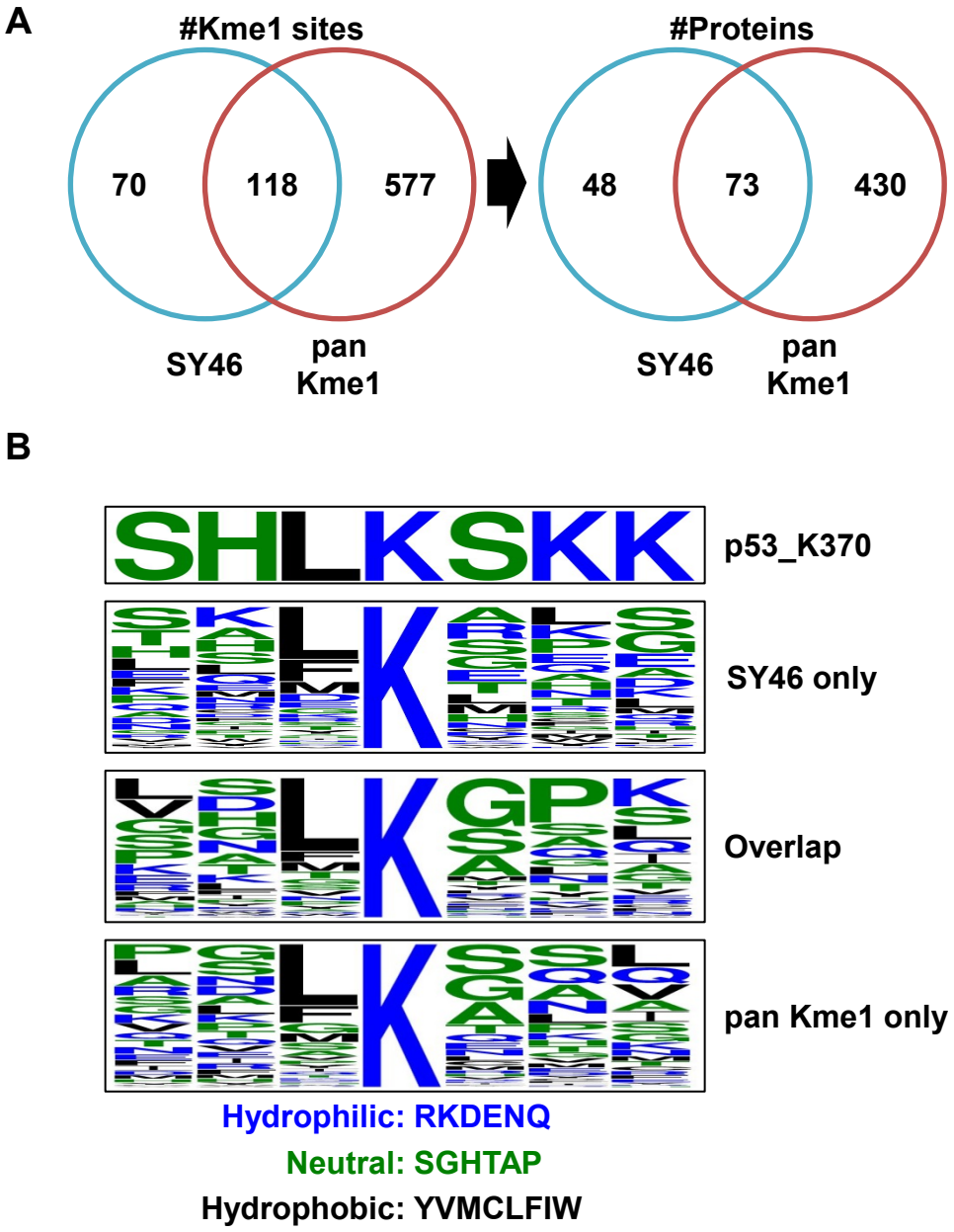


Fig. 4.24: Proteome study reveals novel sites of lysine methylation.
Venn diagrams showing all mono-methylated peptides and corresponding proteins detected either with SY46 antibody or a commercial pan-Kme1 antibody (A). Sequence motifs generated from all detected peptides of the corresponding fraction with the methylated target lysine in the center +/- three neighboring amino acids were compared (B). For comparison, the p53 sequence around lysine 370 is also shown. Different colors represent different hydrophobicity properties of the amino acids according to the WebLogo3 classification [22].

4.2.3 Generation of a candidate list of SMYD2 substrates

Before generating a candidate list of novel SMYD2 substrates, I first checked, if the list of methylated peptides contained already described SMYD2 target sites. Surprisingly, from all so far reported substrates only HSP90 at lysine 607/615 (depending on the isoform) was among the list (**Fig. 4.25**). This might be due to the abundancies of the individual substrates in the MDA-MB231 cell line model. However, the fact that HSP90 methylation was detected, indicated the feasibility of the approach for detecting SMYD2 methylation sites. In order to narrow down the list of potential SMYD2 substrates, I decided to concentrate only on proteins that had been detected by both antibodies, because K607/615 methylation of HSP90 was also in that fraction and the sequence logo of the overlap revealed the strongest enrichment for certain amino acids around the target lysine. In addition, the detection of a methylated peptide in two independent measurements gave a higher confidence that this methylation site can be reproducibly detected. Furthermore, I only concentrated on proteins with a theoretical molecular weight higher than 150 kDa, because of the SMYD2-dependent methylation signal observed on western blots (see previous sections). Using these selection criteria 11 proteins remained (**Tab. 4.9**). When comparing the number of unique methylation sites corresponding to these proteins two proteins were outstanding. Whereas most proteins contained one or two methylation sites, AHNAK and AHNAK2 proteins contained 31 and 10 methylation sites, respectively, that were detected by both antibodies. The high number of methylation sites could explain the strong methylation signals observed on western blots. Therefore, I selected these two proteins as the most promising candidates. However, in subsequent validation experiments I concentrated only on AHNAK protein because AHNAK2 was barely described in the literature and commercial antibodies were only available for AHNAK. In addition, microarray expression data from MDA-MB231 cells suggested that AHNAK expression was much higher in this cell line (**Fig. 4.26**).

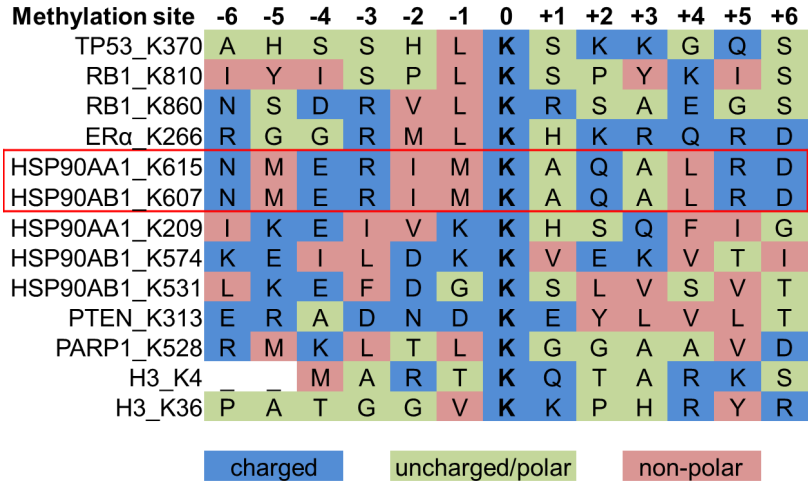


Fig. 4.25: HSP90 K615me1 is the only reported methylation site that is detected in MDA-MB231-SMYD2-WT via mass spectrometry.
Alignment of the reported SMYD2 substrate methylation sites with indicated properties of amino acids. HSP90 at K607/6015 was the only methylation site detected in the present proteome study.

Tab. 4.9: List of genes with proteins >150 kDa and corresponding Kme1 sites

Gene	Mass, kDa	#Kme1 sites SY46	#Kme1 sites pan-Kme1	both
SYNE2	796	1	1	1
AHNAK	**629**	**33**	**77**	**31**
AHNAK2	**617**	**10**	**19**	**10**
PRRC2C	317	2	4	2
GCN1L1	293	1	2	1
MAST4	284	2	2	2
FLNB	278	1	2	1
TNKS1BP1	182	2	3	2
API5	182	1	1	1
GLI2	168	2	2	2
WHSC1	152	2	1	1

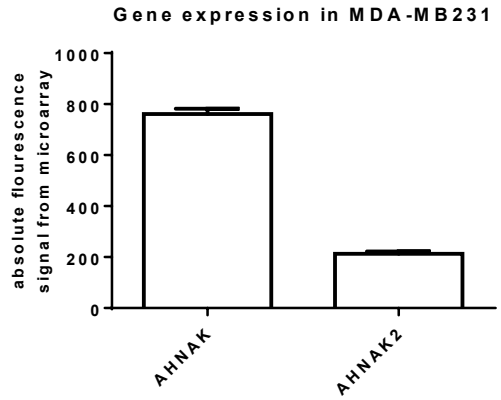


Fig. 4.26: AHNAK and AHNAK2 expression in MDA-MB231.
Comparison of AHNAK and AHNAK2 gene expression in MDA-MB231 from previous gene expression study. The absolute fluorescence signals from microarray suggest much stronger expression of AHNAK in the MDA-MB231 cell line. Error bars indicate standard deviation.

4.2.4 AHNAK is the main methylation substrate detected by SY46

To validate AHNAK as a potential substrate of SMYD2 methylation, I first checked if the methylation signal derived from SY46 corresponded with the localization of AHNAK protein. To this end, MDA-MB231 cells stably overexpressing SMYD2-WT were analyzed by immuno-fluorescence microscopy and co-stained with SY46 (rabbit antibody) and AHNAK (mouse antibody). AHNAK and the SY46 derived methylation signal showed a clear co-localization mainly in the cytosol **Fig. 4.27**. This indicated that the main methylation signal might be derived from AHNAK protein.

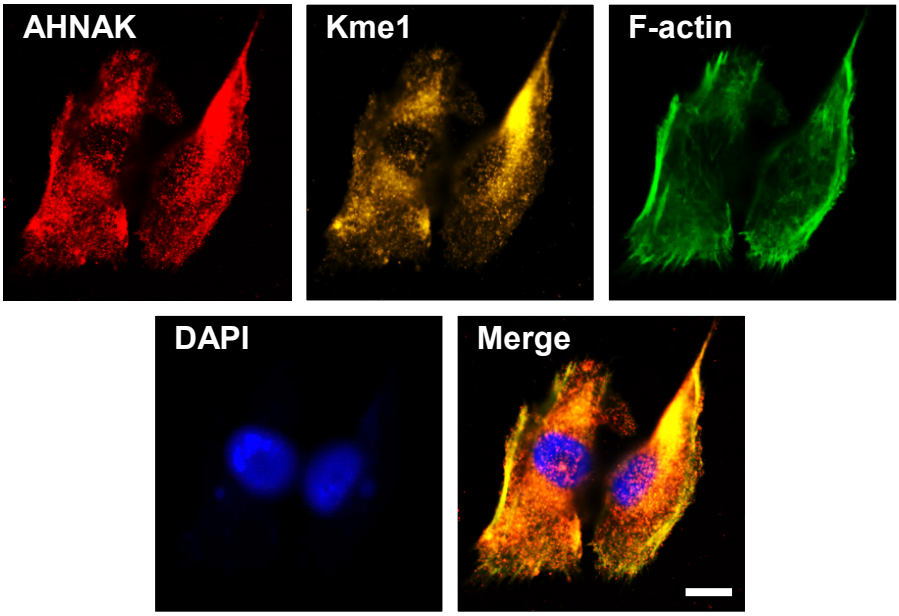


Fig. 4.27: AHNAK co-localizes with SY46 methylation signal.
Immuno-fluorescence staining of stable SMYD2-WT overexpressing MDA-MB231 cells stained for AHNAK, Kme1 (SY46), F-actin (phalloidin) and nuclei (DAPI). Scale bar indicates 20 µm.

To further pursue this question, MDA-MB231 cells overexpressing either SMYD2-WT or the inactive mutant SMYD2-H207A were also analyzed by western blot and detected with SY46 antibody for methylation and for AHNAK. AHNAK protein was detected in both samples at similar levels at the top of the western blot. Strikingly, the methylation signal detected by SY46 overlapped with AHNAK signal in the SMYD2-WT overexpressing sample (**Fig. 4.28A**). This reinforced the above observation from immune-flourescence stainings that co-localization of AHNAK and SY46 derived methylation might be due to AHNAK methylation itself. To finally prove this hypothesis, siRNA mediated knockdown of AHNAK not only diminished AHNAK protein level but also reduced the methylation signal to the same extend (**Fig. 4.28B**). This clearly demonstrated that AHNAK is the main methylation substrate in MDA-MB231 cells that was detected by SY46 antibody.

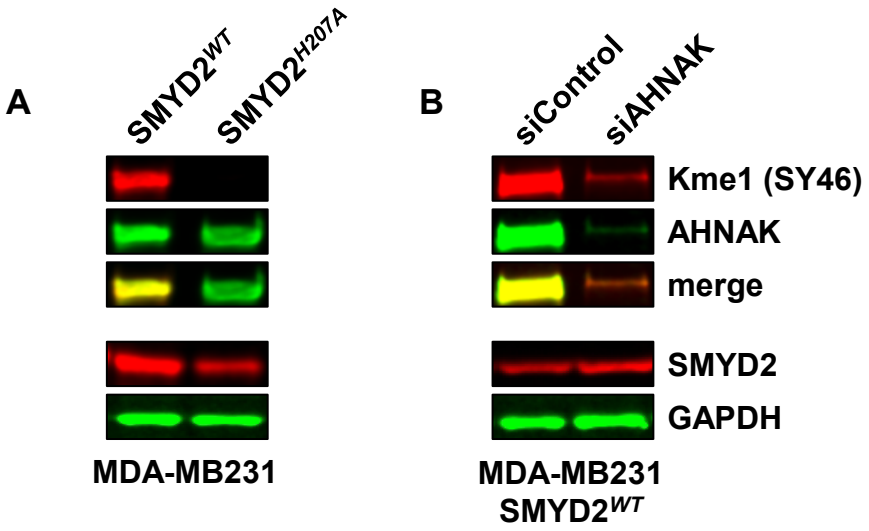


Fig. 4.28: AHNAK is a novel substrate of SMYD2.
Western blot analysis of MDA-MB231 cells stably overexpressing either SMYD2-WT or an enzymatically dead point-mutant SMYD2-H207A. The methylation signal derived from SY46 antibody perfectly overlaps with AHNAK signal (A) and upon AHNAK knockdown the methylation signal decreased to the same extend as AHNAK protein level (B).

4.2.5 Immunoprecipitation of AHNAK enriches also for methylation

In order to check, if enrichment of AHNAK can also enrich for methylation and if SMYD2 and other proteins like HSP90 may form stable complexes with AHNAK dependent on the methylation status, AHNAK protein was immunoprecipitated in MDA-MB231 cells stably overexpressing SMYD2-WT after treatment for 3 days with 1 µM BAY-598 or DMSO (**Fig. 4.29A**). Eluates of AHNAK pulldown showed a clear enrichment of AHNAK protein independent of the treatment. AHNAK methylation was also enriched in the untreated group, but as expected was abolished by BAY-598 treatment. The similar amounts of AHNAK in input and after enrichment for AHNAK in both treated and untreated samples suggested that AHNAK protein stability is not affected by its methylation status. Similarly to AHNAK precipitation, precipitation with SY46 antibody enriched for the methylation signal and for AHNAK in the untreated sample but not in the BAY-598 treated sample (**Fig. 4.29B**). In both experiments no co-purification of SMYD2 or HSP90 was observed in the eluates, indicating only short interactions between SMYD2 and AHNAK during methylation and either weak or no complex formation of AHNAK with HSP90. Interestingly, pull-down with SY46 was also able to specifically capture p53 in the untreated group but not HSP90. This is in contrast to the above proteome study, where methylated peptides were detected from HSP90 but not from p53. An explanation could be that SY46 is not able to capture methylated native full-length HSP90 in contrast to unstructured peptides. The weak signal of p53 methylation after IP may be the reason why it was not detected during the mass spectrometry measurement. It also indicated that only a small fraction of the total p53 pool may actually be methylated even though SMYD2 was ectopically overexpressed.

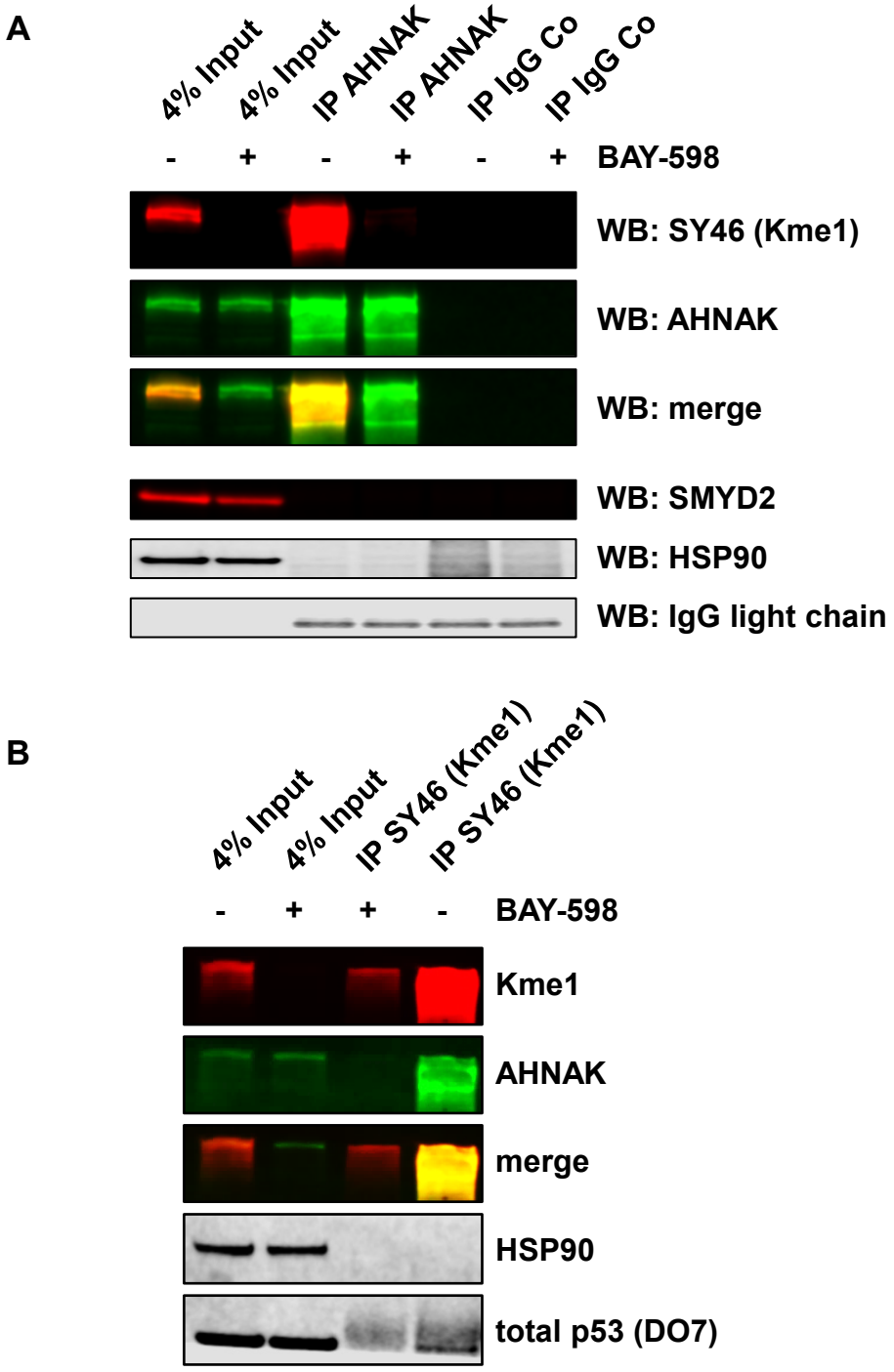


Fig. 4.29: Methylated AHNAK can be immuno-precipitated.
Western blot analysis of stable SMYD2-WT overexpressing MDA-MB231 cells after immuno-precipitation with the indicated antibodies. Cells were treated for 72 h with 1 µM BAY-598 or DMSO.

4.2.6 IC50 of BAY-598 based on AHNAK methylation

As cellular IC50 of BAY-598 was determined by the above mentioned in-cell-western assay of engineered KYSE150 cells (see **3.27**), I was wondering how well this assay correlated with the methylation of AHNAK protein in MDA-MB231 cells. Therefore, MDA-MB213 cells stably overexpressing SMYD2-WT were treated with increasing amounts of BAY-598 for 72 h. AHNAK methylation was determined on western blot (**Fig. 4.30**). Quantification of the methylation signal relative to total AHNAK protein revealed a dose-response curve with an IC50 of BAY-598 of around 50 nM. This is very similar to the cellular IC50 determined by ICW in KYSE150 cells and showed that the cellular IC50 for SMYD2 substrate methylation of BAY-598 is comparable between cell lines.

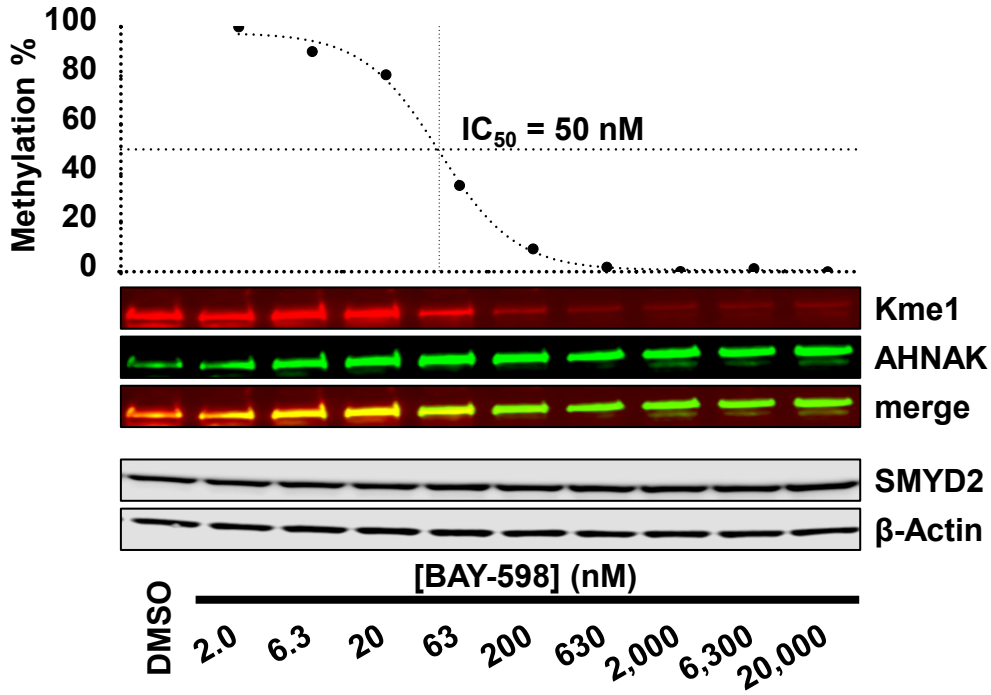


Fig. 4.30: Cellular IC50 of BAY-598 for AHNAK methylation is similar to ICW assay. IC50 determination of BAY-598 based on western blot analysis of AHNAK methylation (SY46). Stable SMYD2-WT overexpressing MDA-MB231 cells were treated with a half-logarithmic dilution series of BAY-598 for 72 h before analysis. For IC50 calculation methylation signals were normalized to AHNAK.

4.2.7 Domain structure of AHNAK and mapping of methylation sites

Human AHNAK protein has been described as a giant scaffold protein of around 623 kDa. The protein was described to be comprised mainly of three parts. First, the N-terminal domain of 498 amino acids, which contains a PDZ domain known as a protein-protein interaction domain. Second, it has a middle part, which consists of ˜38 repeating units referred to as central repeated units (CRUs). These are either ˜128 aa (long) or ˜68 aa (short) in length and are very similar to each other in their amino acid sequence. Finally, there is a 1002 aa long C-terminal, which may contain nuclear localization signals. To figure

out which regions of AHNAK might be methylated by SMYD2, all detected methylation sites of AHNAK from the above proteome study were mapped to the full-length protein sequence. This revealed a tendency of preferred lysine residues within the CRUs and parts of the C-terminus (**Fig. 4.31**. A sequence motif derived from all methylated peptides of AHNAK showed a clear enrichment for a [LF]KGPK motif (**Fig. 4.32**), where mainly the first but also the second lysine residue were found to be methylated. This suggested that this motif is a preferred recognition motif for SMYD2 methylation activity.

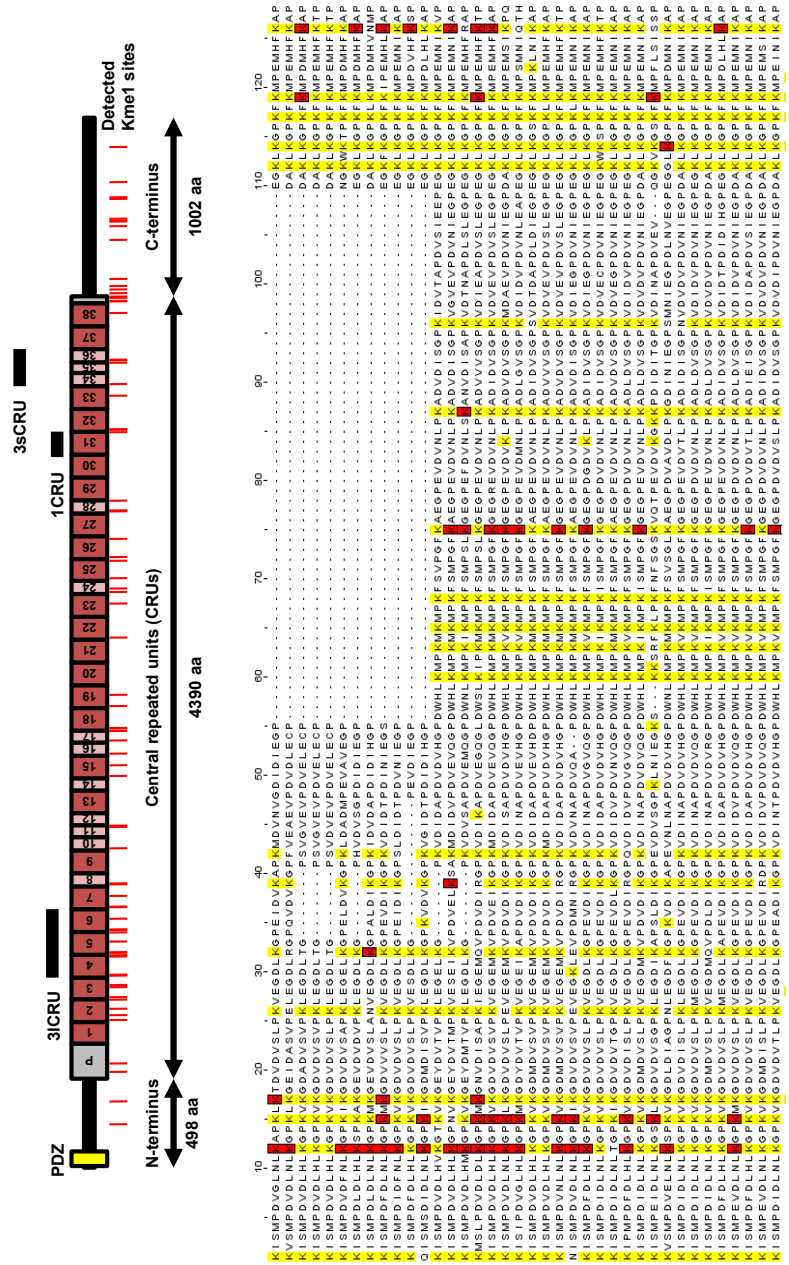


Fig. 4.31: AHNAK domain structure and detected methylation sites.
Top: AHNAK domain structure is illustrated with red dashes indicating methylated lysines detected by mass spectrometry in MDA-MB231 cells ectopically overexpressing SMYD2. The central repeated units (CRUs) have been numbered from 1-38. P means pseudo-repeat (less conserved compared to others), which is less well conserved. PDZ is a protein-protein interaction domain. Recombinant CRUs that have been used for in vitro methylation assays are aligned to the corresponding regions. Bottom: Sequence alignment of the CRUs with short CRUs on top and regular CRUs at bottom. Sequence alignment was adjusted from Kingsley et al.[67]. All lysine residues are marked in yellow. All methylated lysines detected in the proteom study are marked in red.

Fig. 4.32: AHNAK methylation sites reveal enriched sequence motif.
A sequence motif derived from all methylated sites detected on AHNAK protein shows enriched amino acids surrounding the target lysine.

4.2.8 Biochemical SPA methylation assay of SMYD2 using peptides derived from AHNAK

To check, if SMYD2 can directly methylated peptides that contain the enriched motif, an SPA assay, as previously established for p53-peptide to screen for SMYD2 inhibitors, was used. A synthetic peptide derived from AHNAK and containing a lysine, which was detected to be methylated in the proteomics study and containing the LKGPK motif, was synthesized. As control, a p53-peptide was also tested. Unexpectedly, AHNAK peptide completely failed to get methylated in contrast to p53 peptide, which showed increasing methylation over time **Fig. 4.33**. This suggested that either SMYD2 is not able to directly methylate AHNAK peptides without additional cellular components or that peptides were not able to recapitulate structural properties of AHNAK polypeptide, which may be required for the methylation reaction.

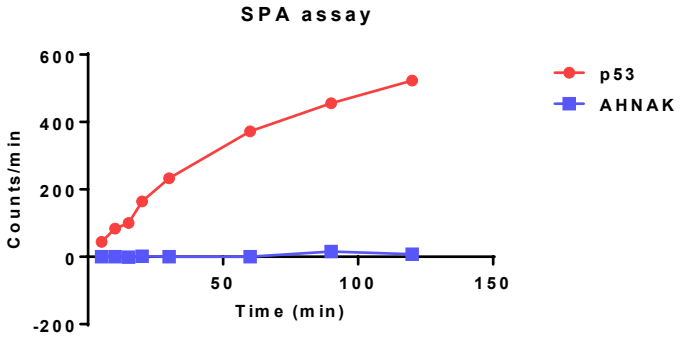


Fig. 4.33: SMYD2 is not able to methylate AHNAK peptides.
The ability of SMYD2 to methylate peptides derived from AHNAK or p53 was tested in a biochemical SPA assay as described in 3.35. Reactions were quenched at indicated time intervals and photon emission was quantified. Peptides derived from AHNAK revealed no signal, whereas p53 peptides were clearly methylated over time.

4.2.9 Design of recombinant AHNAK-CRU fragments for biochemical methylation assays of structured substrate polypeptides

To further investigate, if structured AHNAK polypeptides may serve as a better substrate for SMYD2 in vitro, ideally full-length recombinant AHNAK protein might have been purified and tested in a biochemical methylation assay. However, the large size of this protein made it unsuitable to express the full-length protein by conventional cloning. Therefore, only smaller parts of this protein were cloned into E. coli expression vectors and expressed as GST-fusion proteins. Former reports from literature indicated that it might be possible to express only subunits of the protein and this seemed indeed to be the standard procedure for AHNAK. Therefore, I designed three different constructs that represented different regions of the middle CRU part. All fusion constructs, when aligned to the full-length protein, contained methylation sites that had been detected in the proteome study. These constructs were named 1CRU, 3sCRU (s for short), and 3lCRU (l for long), respectively based on the size and the number of repeating elements (**Fig. 4.31**).

4.2.10 Optimization of AHNAK-CRU expression in E. coli

To obtain sufficient amounts of recombinant GST-fused AHNAK-CRU protein, different expression conditions in E. coli were first tested at small scale. Protein expression was comparable at different conditions but revealed best yields at 0.5 mM IPTG induction and 17°C expression temperature for all three constructs (**Fig. 4.34**). Therefore, this condition was chosen for further protein expression.

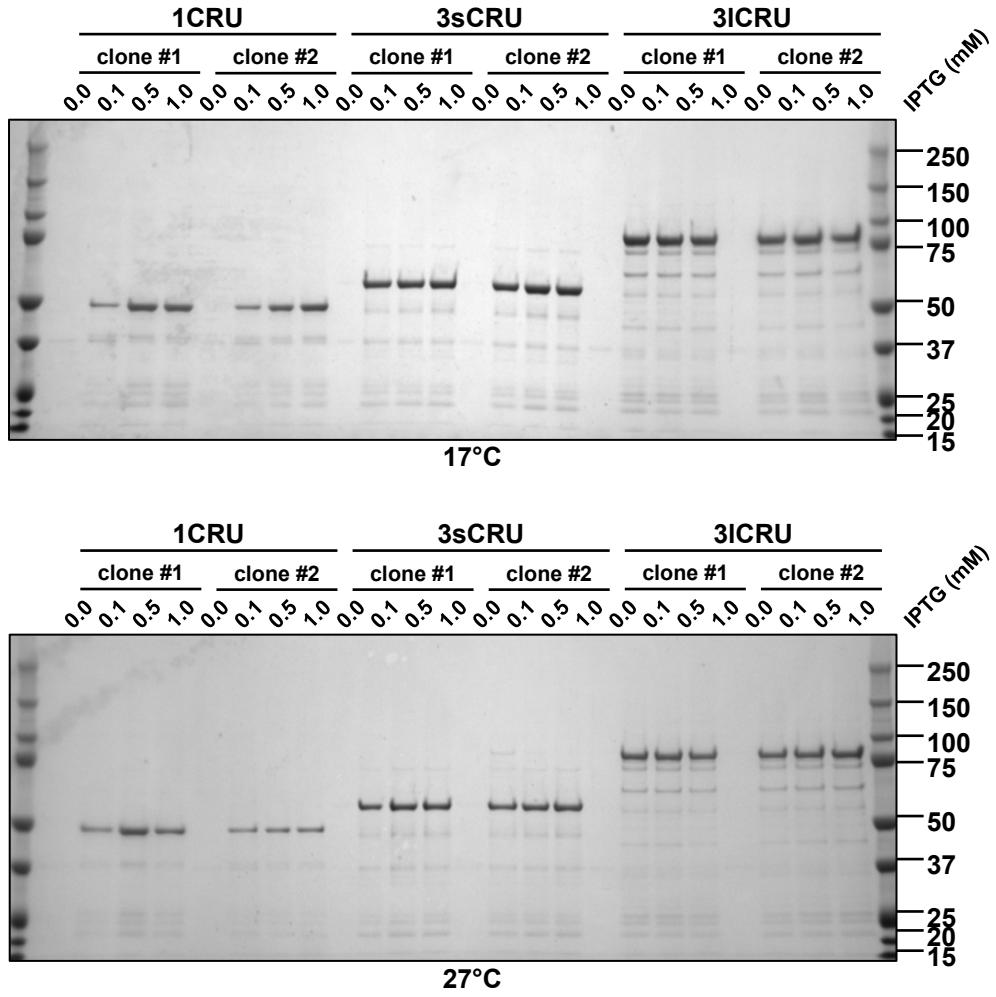


Fig. 4.34: Small scale purification of AHNAK-CRUs for optimization of expression conditions.
Different AHNAK-CRU constructs were transformed in E. coli for recombinant protein expression and cultured under different temperature and IPTG concentrations to reveal optimal expression conditions. All samples were purified via their GST-tags and final eluates were analyzed on a Coomassie-gel. The apparent molecular weight of the prominent protein bands are very similar to the theoretical MWs of 44 kDa, 51 kDa, and 71 kDa for 1CRU, 3sCRU, and 3lCRU, respectively. Numbers on the right indicate molecular weight of the protein markers.

4.2.11 Midi-scale purification of AHNAK-CRUs

For larger amounts of recombinant protein, protein expression was up-scaled to 400 mL E. coli cultures and incubated under optimized conditions. Protein purification was performed via GST-affinity purification (for details see **3.33**). Final eluates revealed clean protein bands corresponding to the calculated MWs of 44 kDa, 51 kDa, and 71 kDa for 1CRU, 3sCRU and 3lCRU, respectively (**Fig. 4.35**).

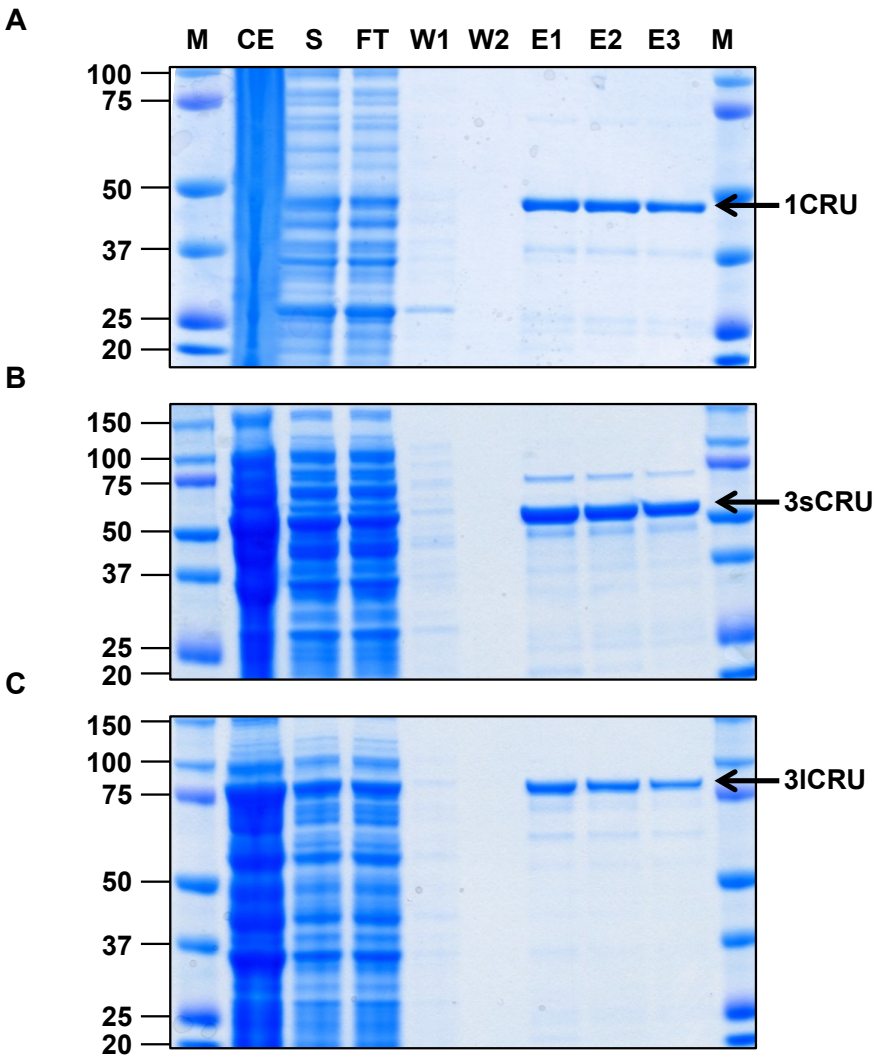


Fig. 4.35: Midi scale purification of AHNAK-CRUs show clear enrichment in the eluate. Different AHNAK-CRU constructs were transformed in E. coli for recombinant protein expression and cultured for 16 h at 17°C, 0.5 mM IPTG. Crude cell lysates were used for the GST purification as described in materials and methods section. Aliquots of each purification step were taken for SDS-PAGE and Coomassie stain. All samples were purified via their GST-tags and final eluates were analyzed on a Coomassie-gel. M – marker, CE – crude extract, S – soluble fraction, FT – flowthrough, W – wash step, E – eluate. Numbers on the left indicate molecular weight of the protein markers.

4.2.12 In vitro methylation of AHNAK-CRUs

To determine, if AHNAK-CRUs can be directly methylated by SMYD2, recombinant GST-fused AHNAK-CRUs (1CRU, 3sCRU, 3lCRU) were incubated +/- SMYD2 enzyme in methylation buffer overnight. Aliquots of the reactions were subjected to SDS-PAGE and analyzed by Coomassie staining and western blot using SY46 methylation specific antibody (**Fig. 4.36**). In fact, SY46 antibody detected clear methylation signals in the samples that contained both AHNAK-CRU substrate and SMYD2 enzyme. In the absence of substrate, weak auto-methylation was observed on SMYD2. These data demonstrate that SMYD2

can directly methylate AHNAK fragments in vitro and that probably more structured AHNAK derived polypeptides are required for efficient methylation. AHNAK-CRUs after in vitro methylation were also cut off from Coomassie gels and analyzed by LC-MS/MS inhouse to detect individual methylation sites. Indeed, several peptides were identified containing a methylated lysines within the LKGPK motif. A representative fragment spectra is shown in **Fig. 4.37**.

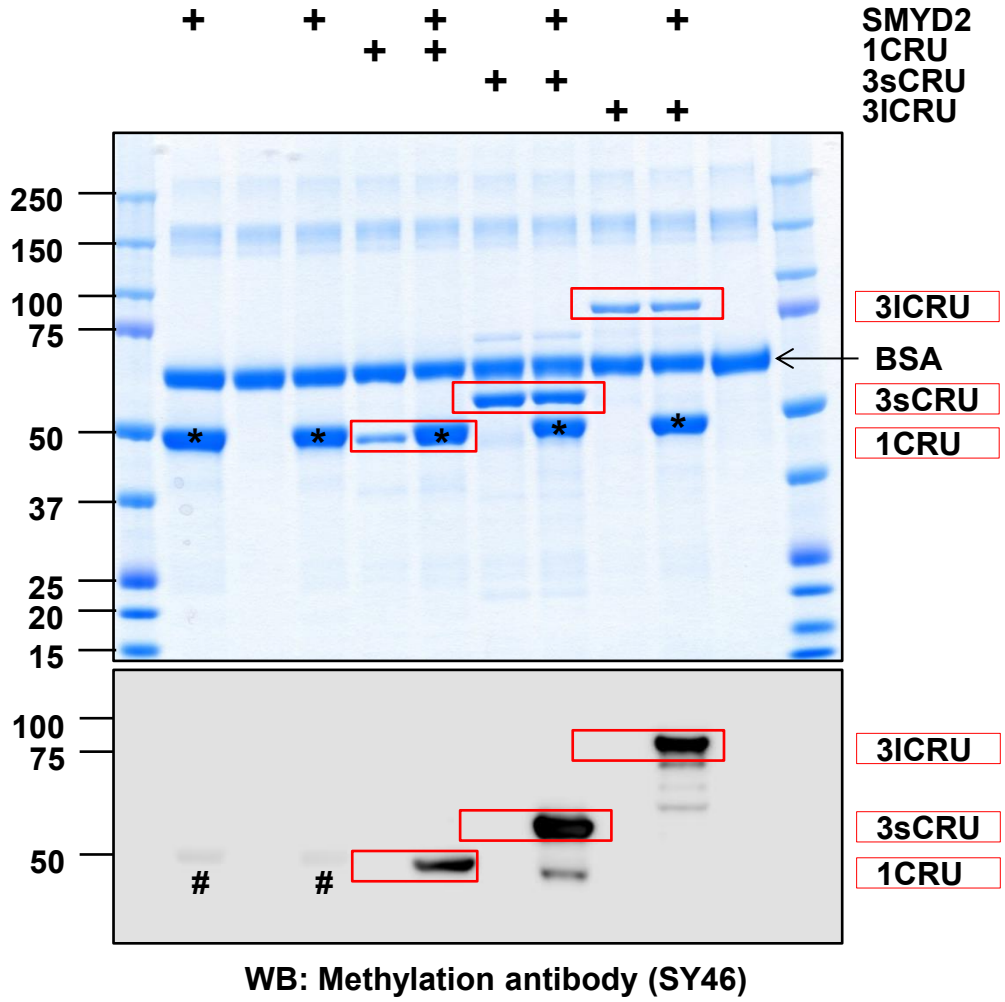


Fig. 4.36: SMYD2 in vitro methylates AHNAK-CRUs.
Coomassie stain (above) and western blot (below) analysis of a biochemical SMYD2 methylation assay using recombinant AHNAK-CRU constructs. The reaction was incubated in methylation buffer overnight at room temperature. Methylation was detected with SY46 antibody. SMYD2 enzyme is marked by asteriks. Hashtag indicate automethylation of SMYD2. Red frames indicate AHNAK-CRU constructs. Numbers on the left indicate molecular weight of the protein markers.

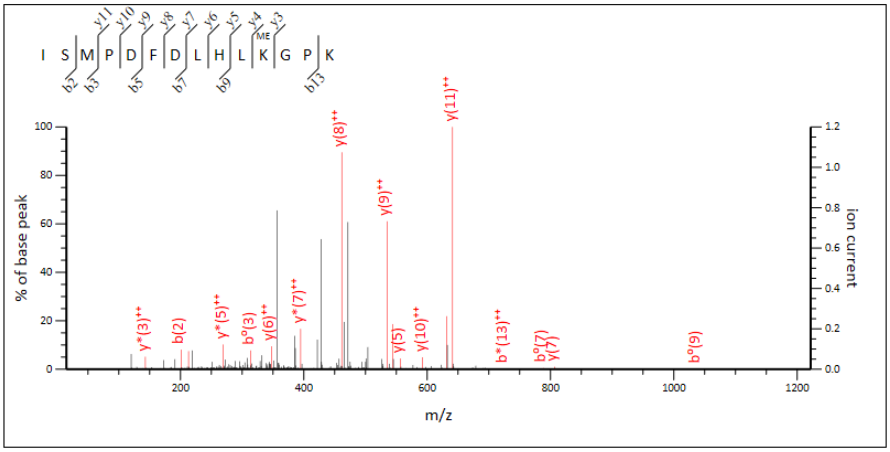


Fig. 4.37: SMYD2 in vitro methylates AHNAK-CRUs at LKGPK motifs.
Exemplary MS/MS spectrum of peptide from AHNAK-3sCRU after in vitro methylation with SMYD2. Monomethylation was detected at lysine within the LKGPK motif. For detection, protein bands from the gel shown in **Fig. 4.36** were cut off and processed for mass spectrometry analysis. Additional fragment spectra are shown in the supplement.

4.2.13 AHNAK methylation occurs in vivo

After demonstrating AHNAK methylation in vitro and in cell lines, I was curious, if AHNAK methylation occurs also in vivo. To test this hypothesis, protein lysates from several mice organs were subjected to immunoprecipitation by SY46 antibody and eluates were analyzed on western blot with AHNAK antibody. Indeed, AHNAK methylation was detected in skeletal muscle tissue and to a weaker extend in cardiac muscle (**Fig. 4.38A**). This correlated very well with the abundance of SMYD2 protein, which was highest in skeletal muscle followed by heart tissue (**Fig. 4.38B**). This observation is consistent with previous reports showing that SMYD2 is highly expressed in these tissues[13]. These data demonstrate that AHNAK methylation occurs in vivo and may possess a relevant physiological function. The observation that AHNAK methylation was detected in human cell lines and in mice also suggests a conserved function for this posttranslational modification in different mammalian species.

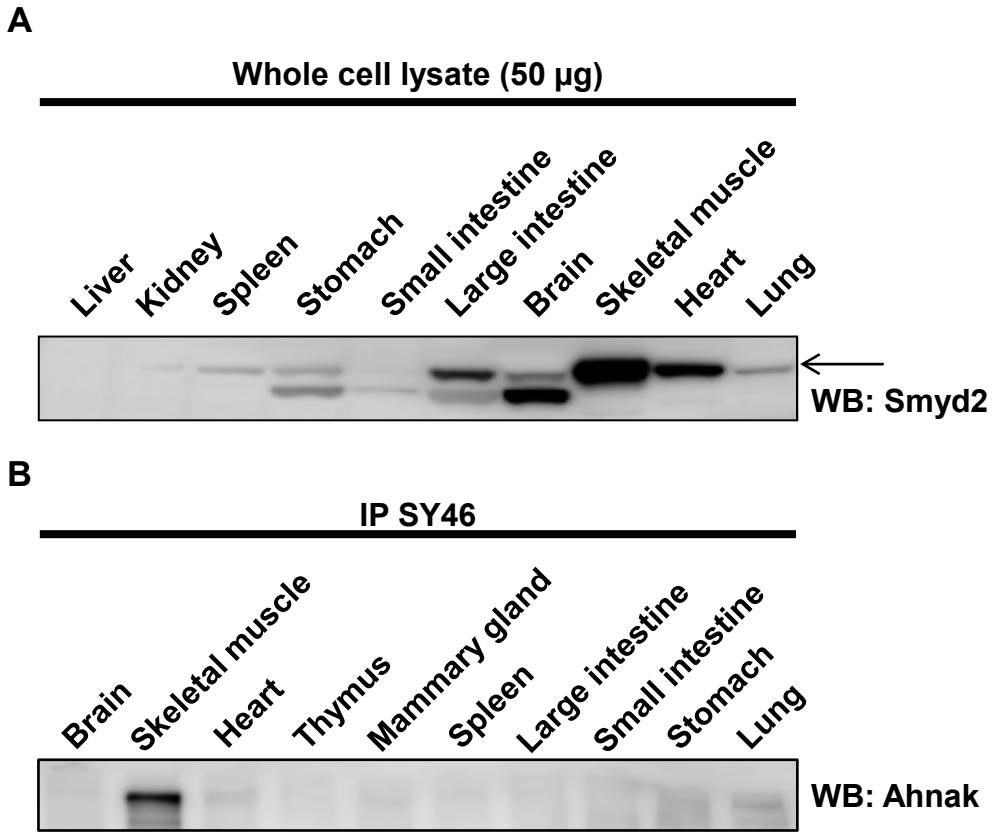


Fig. 4.38: AHNAK is methylated in vivo.
Protein lysates from different mice organs were analyzed for AHNAK methylation and SMYD2 expression. AHNAK methylation was detected by immunoprecipitation with SY46 antibody and detection on western blot with AHNAK antibody.

4.3 Functional characterization of SMYD2 inhibition with the inhibitor BAY-598

After having discovered a novel substrate for SMYD2, and having a potent small molecule probe in hand to specifically inhibit SMYD2 methyltransferase activity, I was interested to test the biological relevance of SMYD2 in selected cellular model systems.

4.3.1 Differentiation of C2C12 myoblasts is not affected by SMYD2 inhibition

Based on my previous findings, which showed that AHNAK methylation occurs in skeletal muscle and that SMYD2 expression is highest in this tissue, I intended to gain a deeper understanding of SMYD2 and AHNAK methylation in muscle biology. For that purpose, I chose the murine myoblast cell line C2C12. Murine C2C12 cells represent an easy to culture cell line model to study muscle cell differentiation and regeneration. The cell line was first established in 1977 by D. Yaffe and O. Saxel to study the functions of homogenous populations of myoblasts[138]. Under subconfluent culture conditions and high levels of fetal calf serum (FCS) these cells will continue to divide as undifferentiated myoblasts. A general differentiation protocol has been established by which cells are grown to confluence and FCS-rich media is changed to serum-low media. This induces fusion of the mycoblasts into myotube like structures – a phenomenon resembeling normal muscle development/regeneration. This process takes several days and involves massive changes in gene expression and phenotype[115]. Therefore, I tested, how SMYD2 might be expresed during that process, and if inhibition of SMYD2 activity by the selective small molecule inhibitor BAY-598 would have an effect during the differentiation process. C2C12 myoblasts were cultured in growth medium (high serum) for three days in the presence of either 1 µM of BAY-598 or DMSO to allow for a sufficient de-methylation phase before starting the actual experiment (start of experiment at day -2, start of differentiation at day 0). Upon differentiation, the treatment groups were kept permanently under BAY-598 (1 µM) and DMSO treatment, respectively. The differentiation process was monitored for 8 days using a light microscope (**Fig. 4.39**) after which myotubes showed spontaneous contractions and started to peel off the cell culture surface (data not shown). During the differentiation process, fusion of individual myoblasts and the progressive increase in myotubes was readily observed. Treatment with BAY-598 did not affect overall cell morphology or myotube formation (**Fig. 4.39**, compare DMSO vs BAY-598). Therefore, I reasoned that inhibition of SMYD2 methylation does not disturb vital functions of C2C12 cells during differentiation.

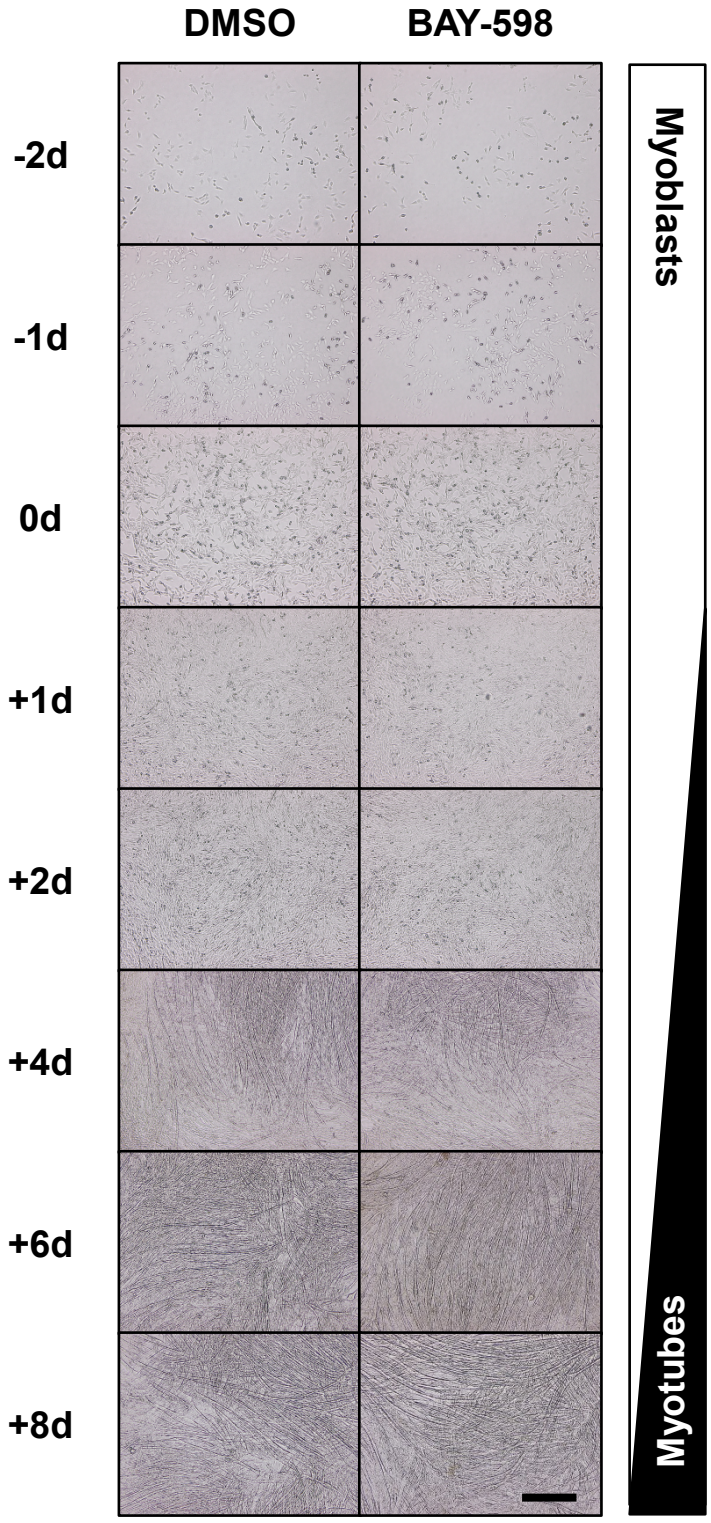


Fig. 4.39: C2C12 differentiation into myotubes is not affected by BAY-598.
C2C12 myoblasts were treated with 1 µM BAY-598 three days prior to initiation of the experiment (experimental start at -2d) to allow a comparison of C2C12 cells harboring completely de-methylated SMYD2 substrates with C2C12 cells of normal SMYD2 substrate methylation state. Cells were kept permanently under BAY-598 treatment to assess the effect of SMYD2 inhibition at all phases of differentiation. Cells were grown in 10 cm dishes and micrographs were taken at different time points as indicated. Scale bar indicates 500 µM.

To characterize gene expression changes of individual genes after SMYD2 inhibition, RNA samples were taken at different time points of C2C12 differentiation and analyzed by qRT-PCR. The expression values of a single gene at different time-points of DMSO and BAY-598 teatment were all normalized to the expression value of that gene from DMSO treatment at time point zero (for details see **3.17**). This allowed for the visualization of changes in gene expression over time and between treatment groups. As differentiation processes in general lead to massive changes in global gene expression patterns, relative gene expression data were not calculated with the conventional ΔΔCt-method, because this method requires inert "house-keeping" genes that are unaffected in its expression rate by the experimental conditions. However, genes like ACTB (encoding for beta-actin), GAPDH (metabolic enzyme), RPL41 (encoding for the 60S ribosomal protein L41) or RPS24 (encoding for the 40S ribosomal protein S24), which are generally considered as house-keepers, showed relatively strong changes in gene expression and were therefore unsuitable as reference genes for the ΔΔCt method (**Fig. 4.40**).

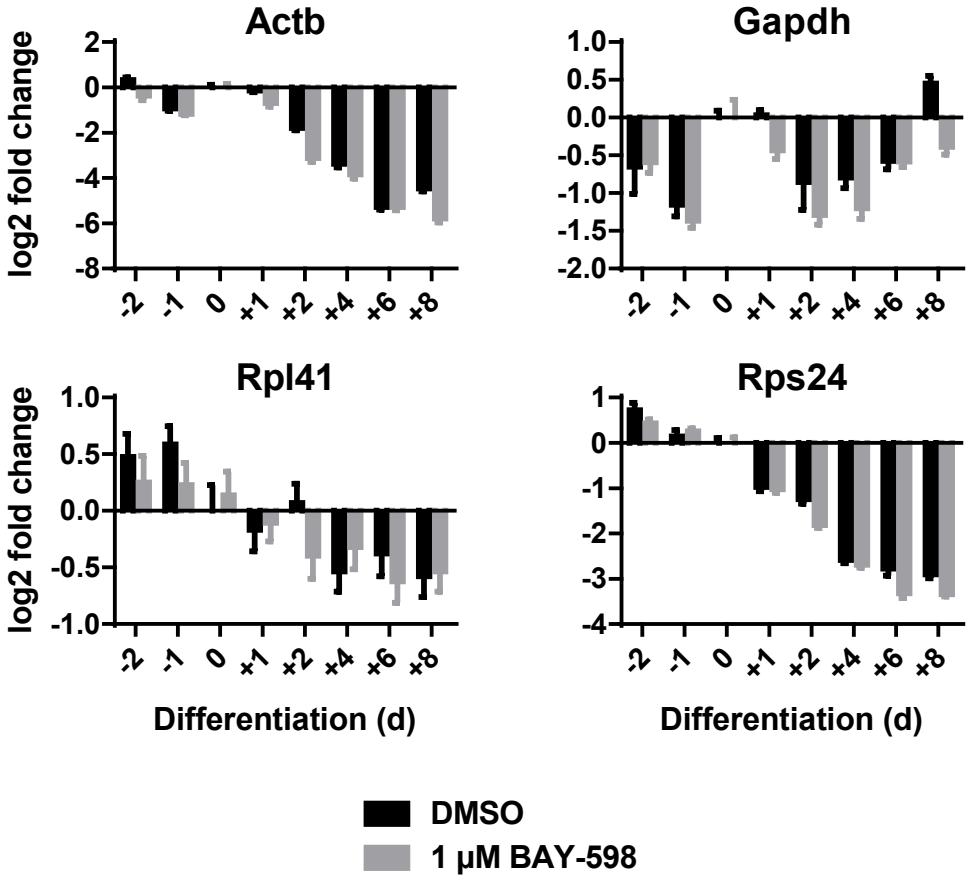


Fig. 4.40: Evaluation of house-keeping genes during C2C12 differentiation.
C2C12 myoblasts were treated with 1 µM BAY-598 three days prior to initiation of the experiment (experimental start at -2d) to allow a comparison of C2C12 cells harboring completely de-methylated SMYD2 substrates with C2C12 cells of normal methylation state. Cells were kept permanently under BAY-598 treatment to assess the effect of SMYD2 inhibition at all phases of differentiation. Cells were grown in 10 cm dishes and RNA samples were prepared for qRT-PCR analyzes. Gene expression data are presented as log2-fold changes relative to time point zero of DMSO treatment.

Therefore, I decided to deploy the Δ'Ct method[83], which normalizes the Ct values to the total input of RNA/cDNA template. This method seemed appropriate as it showed expected changes in gene expression of several biomarker genes of muscle differentiation. For instance, genes like DMD (encodes for dystrophin, a vital component of the cytoskeleton in muscle fibers;[40]), MEF2C (a known transcription factor of the MEF2-family involved differentiation and myogenesis[34]), and MYOG (a vertebrate specific transcription factor known for its role in muscle differentiation and regeneration[102]) served as positive biomarkers of myotube differentiation and showed a striking induction upon initiation of differentiation of ~8-fold (MEF2C) and up to ~1000-fold (DMD, MYOG) (**Fig. 4.41**). In contrast, the transcription factor MYOD was not significantly regulated upon induction and showed even a slight down-regulation (**Fig. 4.41**). This might be explained by the fact, that MYOD is one of the earliest markers of myogenic commitment and only silent in quiescent satellite cells but not in proliferating myoblasts[102]. In all cases, treatment with BAY-598 had no significant influence on these biomarkers, which is in agreement with the undisturbed morphological changes observed above on micrographs (**Fig. 4.39**).

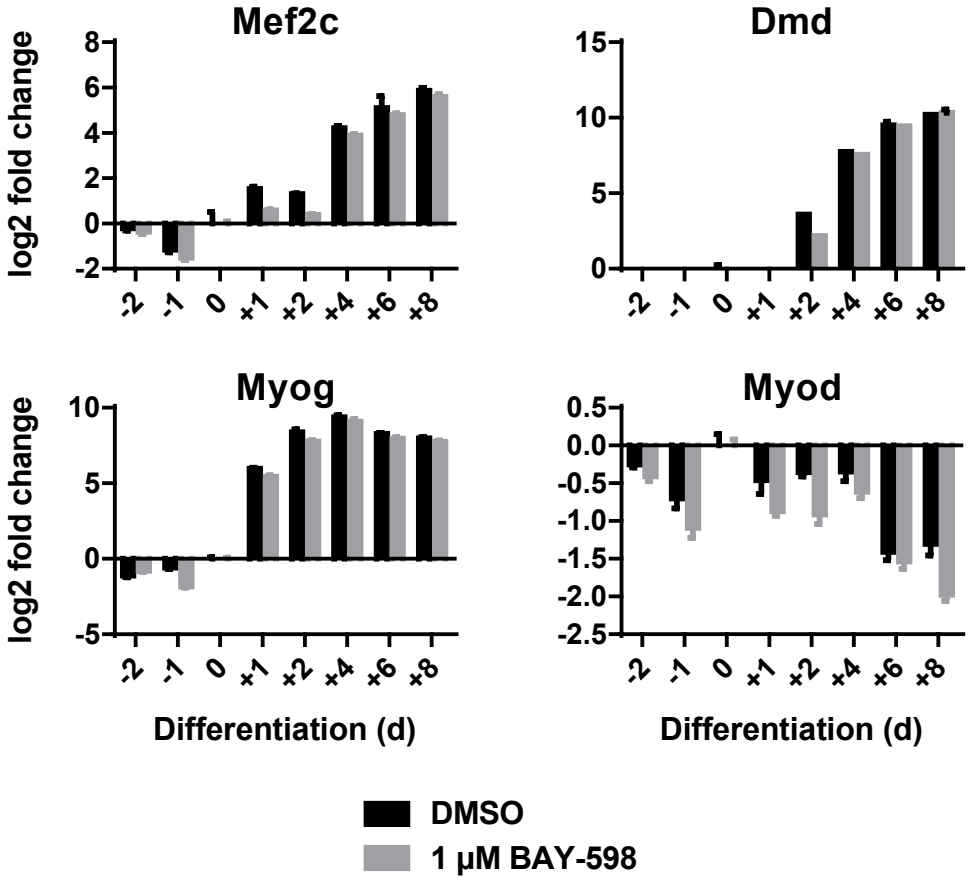


Fig. 4.41: C2C12 biomarker genes indicated proper myotube differentiation.
C2C12 myoblasts were treated with 1 µM BAY-598 three days prior to initiation of the experiment (experiment start at -2d) to allow a comparison of C2C12 cells harboring completely demethylated SMYD2 substrates with C2C12 cells of normal methylation state. Cells were kept permanently under BAY-598 treatment to assess the effect of SMYD2 inhibition at all phases of differentiation. Cells were grown in 10 cm dishes and RNA samples were prepared for qRT-PCR analyzes. Gene expression data are presented as log2-fold changes relative to time point zero of DMSO treatment. n.d. = not detected.

In order to evaluate, if other SMYD-family members may compensate for SMYD2 inhibition, I compared the expression of all SMYD family members during the differentiation process. Interestingly, different patterns of regulation were observed between SMYD genes (**Fig. 4.42**). SMYD2, SMYD4, and SMYD5 were expressed at constant levels during the myoblast phase (-2d – 0d), but started to continuously decrease in their expression during the differentiation process, which reached a maximum of down-regulation of about 4-8 fold after 6 days of differentiation compared to day zero. The pattern of gene expression changes among these three genes may indicate a common transcriptional regulation and a higher relevance of these genes in undifferentiated myoblasts. SMYD3 expression was more fluctuating and did not show a clear positive or negative correlation with the differentiation status. In contrast, SMYD1 behaved in the opposite direction compared to SMYD2, SMYD4, and SMYD5 and showed a strong up-regulation after induction of differentiation, which reached a plateau of around 64-fold increase in transcript level after four days of differentiation compared to day zero. This is in alignment to reports, showing that SMYD1 is highly expressed in striated muscle tissue and is important for myofilament organization[41;90]. In any case, treatment with BAY-598 did not show a significant impact on transcript levels of either SMYD2 nor the other SMYD-family members, suggesting that SMYD2 inhibition is not compensated by transcriptional up-regulation of other family members.

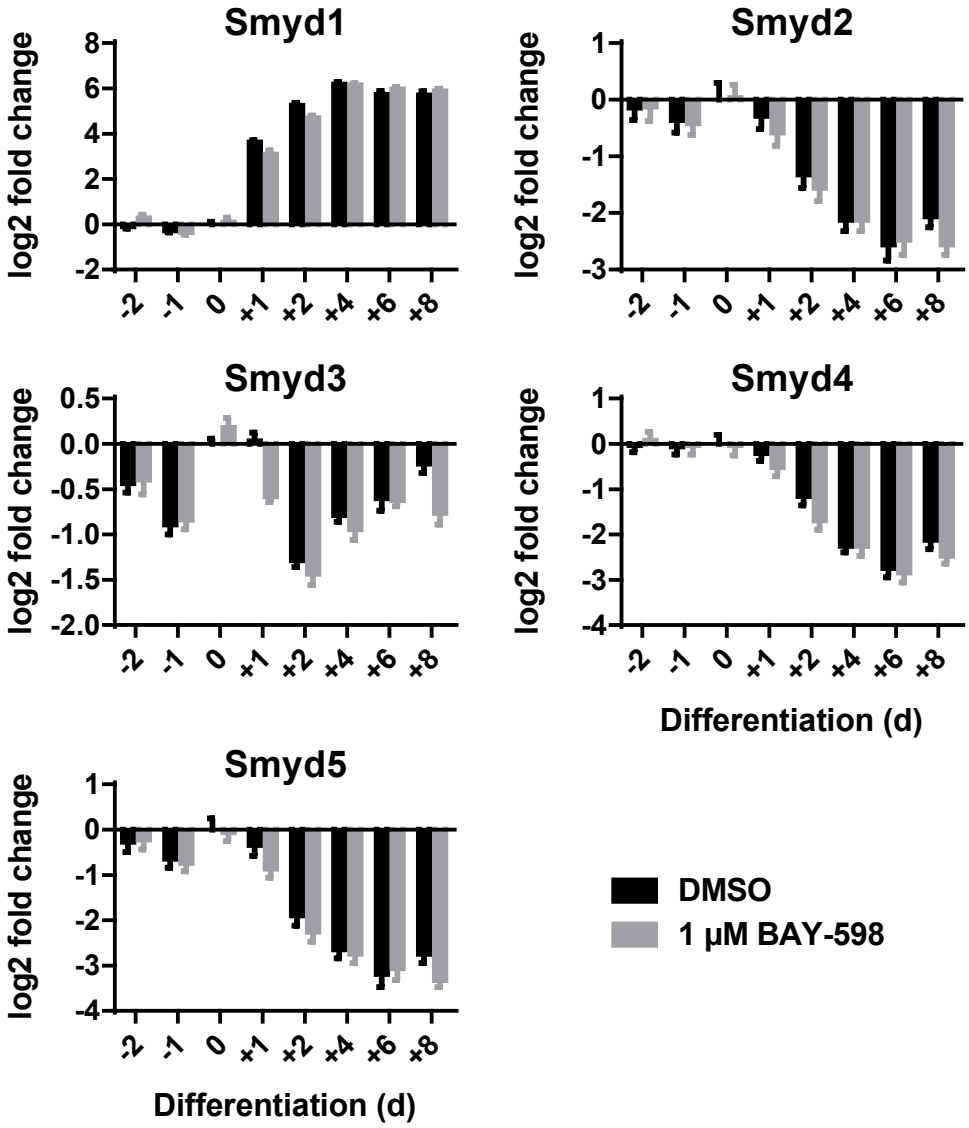


Fig. 4.42: Other SMYD-family members seem not to compensate for SMYD2 inhibition. C2C12 myoblasts were treated with 1 µM BAY-598 three days prior to initiation of the experiment (experimental start at -2d) to allow a comparison of C2C12 cells harboring completely de-methylated SMYD2 substrates with C2C12 cells of normal methylation state. Cells were kept permanently under BAY-598 treatment to assess the effect of SMYD2 inhibition at all phases of differentiation. Cells were grown in 10 cm dishes and RNA samples were prepared for qRT-PCR analyzes. Gene expression data are presented as log2-fold changes relative to time point zero of DMSO treatment.

Finally, I was interested in the gene regulation of the validated SMYD2 substrate AHNAK and its paralog AHNAK2, which, due to its conservation of the lysine target sites, also represents a likely SMYD2 substrate. Interestingly, both genes were reversely regulated (**Fig. 4.43**). AHNAK expression was induced over ten-fold during differentiation compared to the undifferentiated myoblast state (days -2 and -1) whereas AHNAK2 expression was decreased up to around ten-fold during differentiation. Of note, absolute Ct values suggested, that both genes were expressed at similar degrees (median Ct of AHNAK changed from

26.4 at -2d to 23.2 at 8d, whereas median Ct from AHNAK2 changed from 23.4 at -2d to 26.4 at 8d). Comparing this pattern to the SMYD2 expression pattern would suggest that in this model system probably AHNAK2 is more likely to be a relevant substrate for SMYD2. However, SMYD2 inhibition by BAY-598 treatment also revealed no changes in gene expression of its potential substrates AHNAK and AHNAK2. Taken together, these data indicate that SMYD2 inhibition does not induce a feedback loop, which would change the transcriptional activity of itself, its family members, or its substrates AHNAK and AHNAK2, et least under the tested conditions.

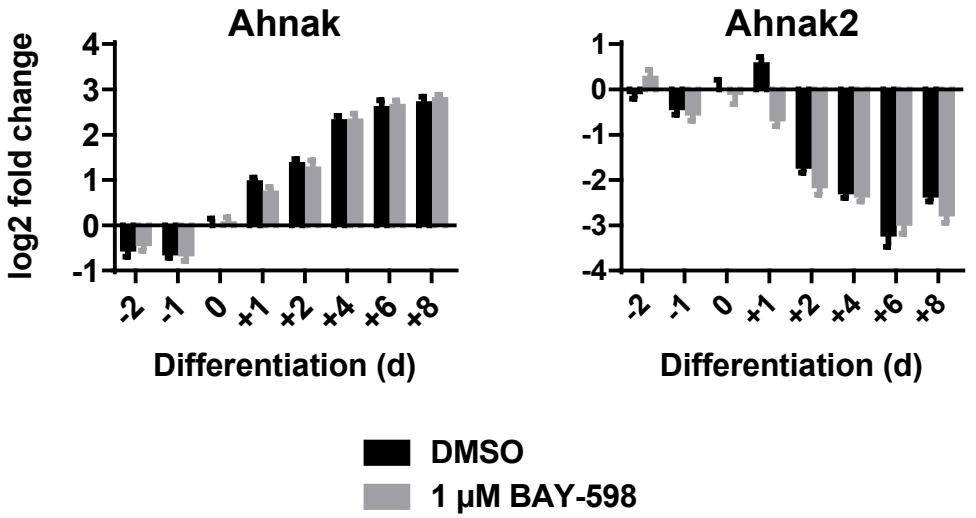


Fig. 4.43: Gene expression patterns suggest Ahnak2 as dominant substrate SMYD2 in C2C12.
C2C12 myoblasts were treated with 1 µM BAY-598 three days prior to initiation of the experiment (experimental start at -2d) to allow a comparison of C2C12 cells harboring completely de-methylated SMYD2 substrates with C2C12 cells of normal methylation state. Cells were kept permanently under BAY-598 treatment to assess the effect of SMYD2 inhibition at all phases of differentiation. Cells were grown in 10 cm dishes and RNA samples were prepared for qRT-PCR analyzes. Gene expression data are presented as log2-fold changes relative to time point zero of DMSO treatment.

To check, if SMYD2 inhibition may have an impact on the above mentioned genes at the protein level, I performed western blot analyzes of the corresponding protein samples (**Fig. 4.44**). In general, western blot data revealed good correlations between transcript levels and protein levels for SMYD1, SMYD2, and SMYD3. However, after two days of differentiation initiation, stabilization of SMYD2 protein was observed in the BAY-598 treatment group compared to DMSO treatment. This may indicate a compensatory regulation of SMYD2 activity to rescue the substrate demethylation. However, it is also possible that binding of the compound itself to the enzymatic pocket led to a stabilization of the protein, a phenomenon, which has also been observed from other inhibitors.
SMYD5 antibody signal slightly increased during differentiation contrary to RNA transcript levels, which decreased. This may indicate additional levels of regulation beyond transcriptional level. Furthermore, the half-life of SMYD5 potentially might be very long, so that even after decreasing transcript levels the protein accumulates. For SMYD4 no antibody was available. At protein level myogenin and myosin heavy chain (MYHC) served as biomarker controls for differentiated myotubes and showed a clear induction after 1 day

and 4 days of differentiation, respectively. Similar to transcript results, MYOD protein level was present at similar amount at all stages of differentiation.
Unfortunately, I was not able to detect AHNAK protein levels in a consistent way for all samples, as the protein tended to enrich in the non-dissolved cell debris of the protein lysates with increasing differentiation and the attempt to dissolve cell debris resulted in unequal degradation of the protein between samples, which made it unfeasible to compare protein levels of different samples (data not shown). Therefore, also AHNAK methylation could not be adequately determined. As a summary of this differentiation experiment I concluded that SMYD2 expression is higher in more undifferentiated myoblasts and according to transcript levels that AHNAK2 may present a more likely substrate in this system. However, SMYD2 inhibition did not seem to influence myotube differentiation. This finding supports an earlier study, where SMYD2 knockout mice showed no developmental defect and were undiscernible from its wildtype littermates at least under normal conditions[103].

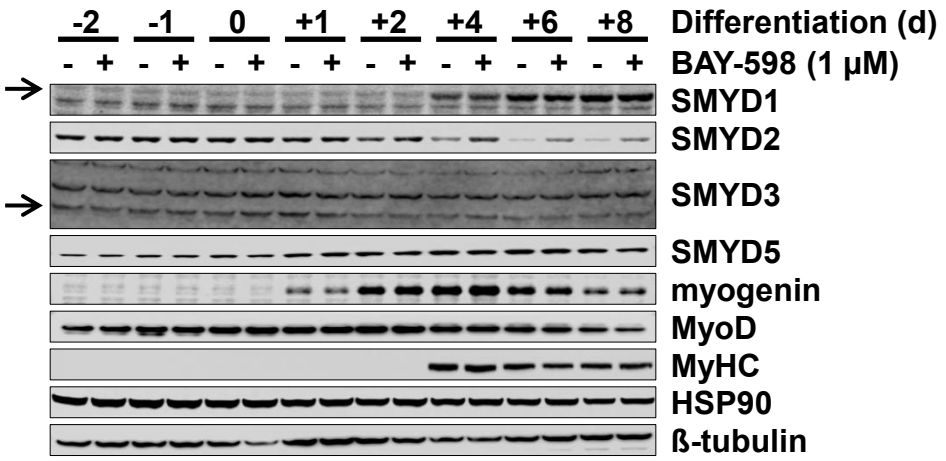


Fig. 4.44: SMYD2 protein expression decreases in differentiating C2C12.
C2C12 myoblasts were treated with 1 µM BAY-598 three days prior to initiation of the experiment (experimental start at -2d) to allow a comparison of C2C12 cells harboring completely de-methylated SMYD2 substrates with C2C12 cells of normal methylation state. Cells were kept permanently under BAY-598 treatment to assess the effect of SMYD2 inhibition at all phases of differentiation. Cells were grown in 10 cm dishes and protein samples were analyzed by western blot.

4.3.2 SMYD2 activity does not impact H-RAS transformation of NIH3T3 cells

In a recent report Reynoird and colleagues presented results showing that adenocarcinoma formation in the pancreas induced by oncogenic RAS transformation led to a strong upregulation of SMYD2 protein expression, which is normally not detectable in healthy pancreas. Furthermore, RAS transformation was significantly compromised by conditional knockout of SMYD2 in the pancreatic cells[103]. They further provided mechanistic evidence that SMYD2 methylates a MAP kinase involved in the RAS-ERK1/2 pathway and thereby promoting ERK1/2 phosphorylation and oncogenesis.

Based on this study, I was curious to test whether the observed effects can be recapitulated in a simple NIH3T3 cell line model. NIH3T3 mouse fibroblasts have been a useful system to study the transformation potential of many known oncogenes, including all RAS variants.

SMYD2 compound inhibition during focus formation I employed a classical NIH3T3 focus formation assay by introducing oncogenic H-Ras-G12V[99] via retrovirus transduction (mutation from glycine to valine at amino acid residue 12 keeps RAS permanently activated by preventing GTP hydrolysis to GDP, which normally reverts RAS activation). Transduced cells were permanently treated with DMSO as control, 1 µM of the SMYD2 inhibitor BAY-598 or 100 nM of Refametinib (biochemical IC50 for MEK1 = 19 nM, for MEK2 = 47 nM), a dual MEK1/2 inhibitor, which is able to block the RAS/RAF/MEK/ERK pathway of oncogenic RAS signaling[110]. Transduced cells were also compared to the non-transduced untreated parental cells. The focus formation assay was evaluated ten days after transduction. Acidification of the cell culture media and changes in cell morphology clearly indicated effects of oncogenic Ras activity in all transduced cells compared to the parental cells (**Fig. 4.45**, light microscope). Parental cells showed dense monolayers, whereas Ras transduced cells were radial shaped with intercellular space between adjacent cells and appeared to be more shiny under a phase contrast light microscope. However, whereas the MEK1/2 inhibitor Refametinib was able to prevent focus formation, clear focus formation was observed in both DMSO and BAY-598 treated cells. This suggested that SMYD2 enzymatic inhibition was not able to prevent focus formation.

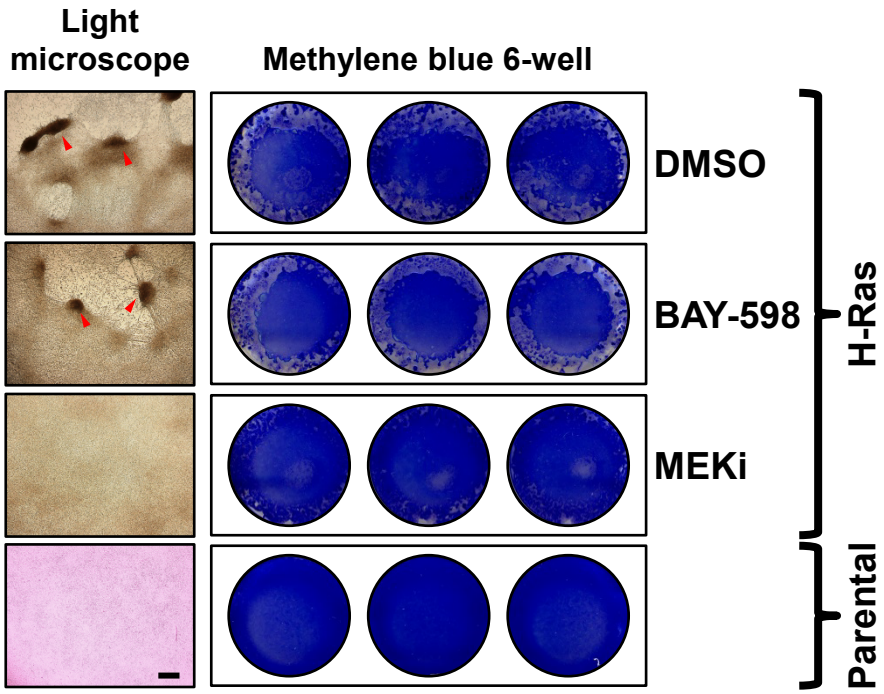


Fig. 4.45: BAY-598 does not prevent focus formation in oncogenic RAS transduced NIH3T3 cells.
NIH3T3 mouse fibroblasts were seeded in triplicates in 6-well plates and transduced with retrovirus encoding an oncogenic H-RAS (G12V) variant. Cells were cultured either in the presents of DMSO control, 1 µM BAY-598 or 100 nM of the dual MEK1/2 inhibitor Refametinib for up to ten days. Foci were visualized either by light microscopy (left) or methylene blue staining. Scale bar indicates 500 µM.

Genetic SMYD2 knockout during focus formation As effects from enzymatic inhibition using small molecule inhibitors can significantly differ from genetic ablation of the whole protein especially when active in protein complexes, I established NIH3T3-CrispR/Cas9 SMYD2 knockout (SMYD2-KO) cells. Immuno-fluorescence analyzes revealed a successful knockout and a solely cytosolic localization of endogenous SMYD2 in the parental cells (**Fig. 4.46**).

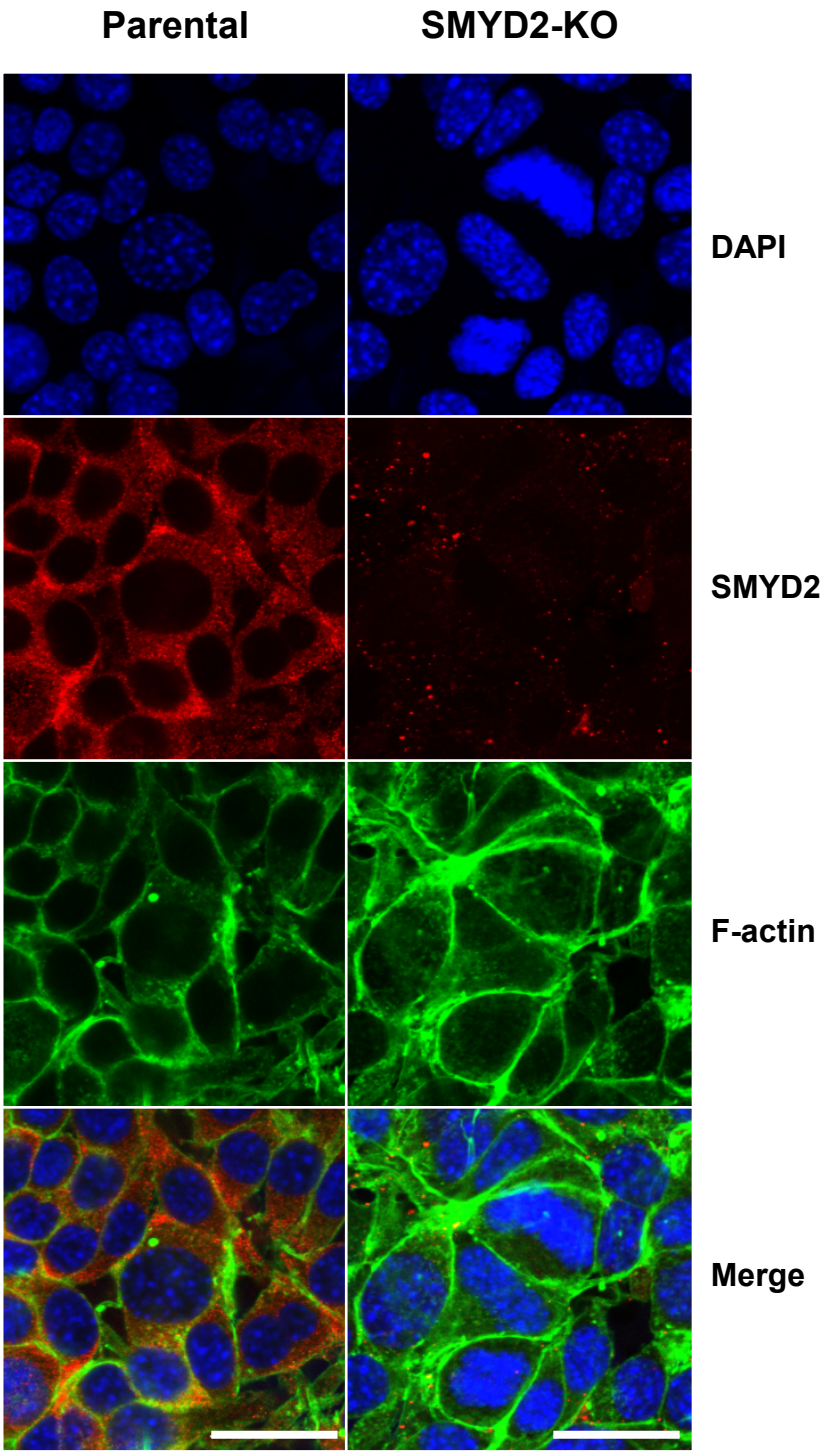


Fig. 4.46: SMYD2 localizes solely to the cytoplasm in NIH3T3 fibroblasts
NIH3T3 mouse fibroblasts either parental or a CrispR/Cas9 SMYD2 gene knockout were analyzed by immuno-fluorescence for SMYD2 localization. DAPI was used to stain nuclei and F-actin was visualized by Alexaflour 555-phalloidin. Scale bar indicates 50 µm.

Having this SMYD2-KO cell line generated, a focus formation assay in combination with RAS-G12V transformation as above was repeated. SMYD2-KO cells were also comple-

mented with ectopic SMYD2-WT or GFP as a control. However, RAS transformation revealed focus formation and morphological changes to the same extent in parental cells versus SMYD2-KO cells. In addition, ectopic SMYD2 overexpression did not increase focus formation nor was it able to induce focus formation after 20 days by itself, when transduced without RAS-G12V.

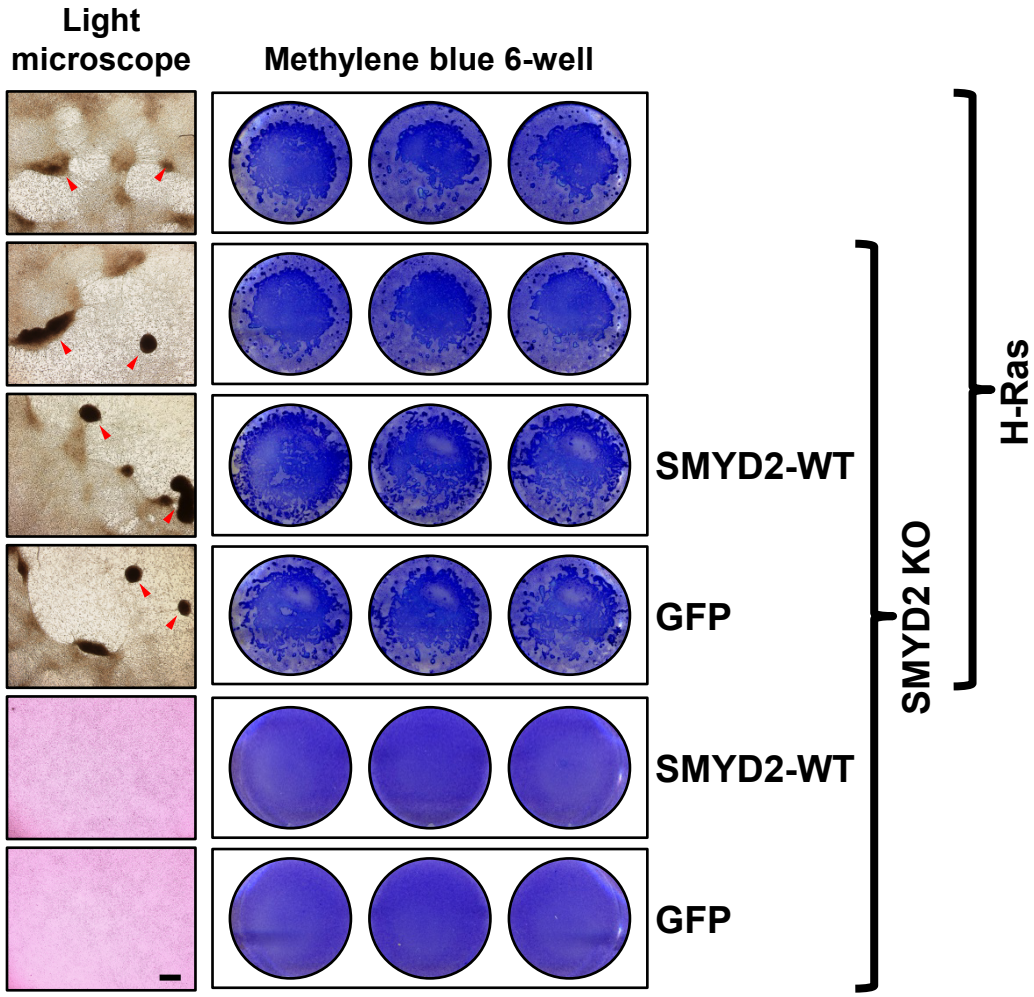


Fig. 4.47: SMYD2 is not an oncogenic enhancer of RAS-G12V transformation in NIH3T3 fibroblasts
NIH3T3 mouse fibroblasts with endogenous levels of SMYD2 or a CrispR/Cas9 SMYD2-KO cell line or SMYD2-KO cells complemented with ectopic SMYD2-WT or GFP were seeded in triplicates in 6-well plates and transduced with retrovirus encoding an oncogenic H-RAS (G12V) variant. Foci were visualized either by light microscopy (left) or methylene blue staining ten days (RAS-transformed) or 20 days (non-RAS transformed) after initiation of the experiment. Scale bar indicates 500 µm.

4.3.3 ERK1/2 phosphorylation is not affected by SMYD2

To also evaluate, if RAS transformation leads to upregulation of SMYD2 expression as reported by Reynoird et al., and if SMYD2 knockout compromises ERK1/2 phosphoryla-

tion, I evaluated protein samples of the above mentioned NIH3T3 cell lines. Indeed, RAS transformation led to an increase in ERK1/2 phosphorylation compared to parental non-transformed cells, but SMYD2 knockout had no influence in the degree of phosphorylation neither in the RAS transformed cells nor in the non-transformed cells (**Fig. 4.48A**). Furthermore, RAS transformed cells showed no upregulation of SMYD2 compared to parental cells. However, immunoprecipitation using SY46 methylation specific antibody revealed endogenous AHNAK methylation in NIH3T3, which was absent in SMYD2-KO cells (**Fig. 4.48A**). In addition, when complementing SMYD2-KO cells either in the background of RAS transformation or in non-transformed cells, SY46 methylation at the size of AHNAK protein was readily observed in whole cell lysates. However, ERK1/2 phosphorylation was not changed upon ectopic SMYD2 overexpression (**Fig. 4.48B**). These data indicated that, in contrast to the pancreatic cancer in vivo model from Reynoids et al., SMYD2 protein abundance in NIH3T3 did not influence ERK1/2 phosphorylation nor did it influence the transformation potential of these cells by oncogenic H-RAS.

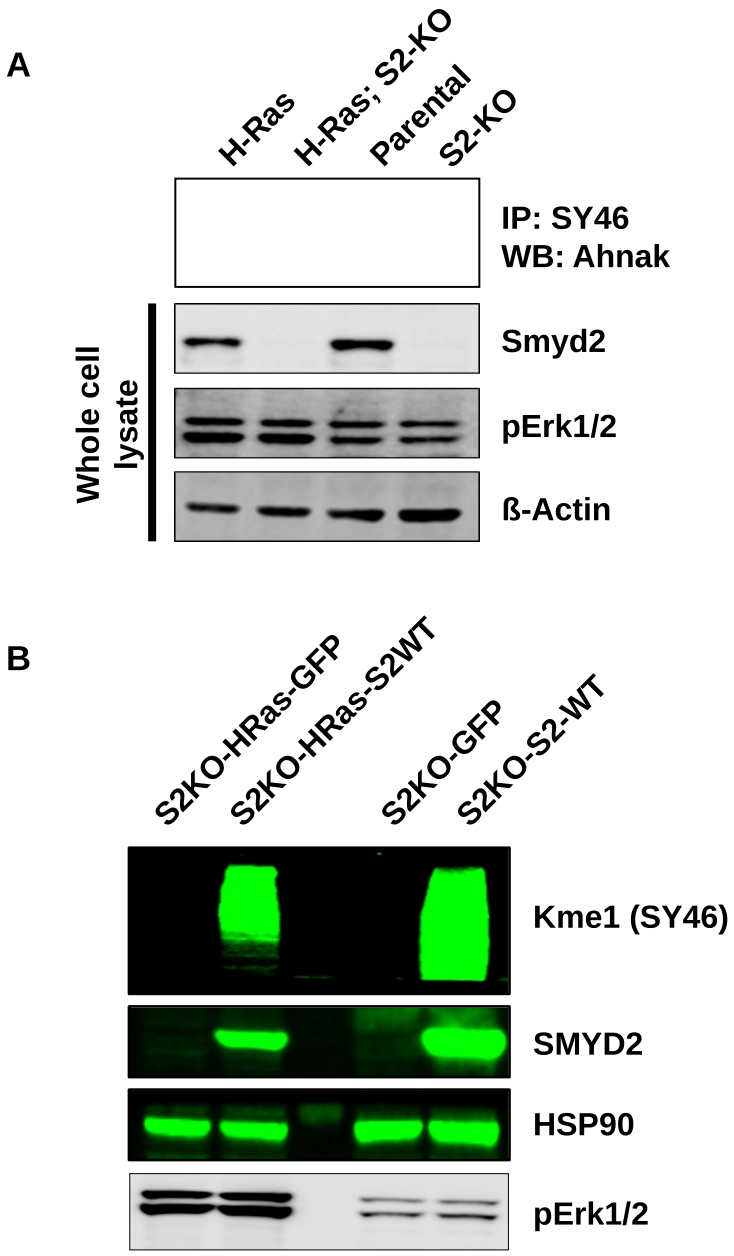


Fig. 4.48: SMYD2 is not an oncogenic enhancer of RAS transformation in NIH3T3 fibroblasts.
(A) NIH3T3 mouse fibroblasts with endogenous levels of SMYD2 (Parental) or a CrispR/Cas9 SMYD2 gene knockout (S2KO) were compared on western blot with cells derived from those, which were additionally transformed with oncogenic H-RAS G12V (HRAS, HRAS-S2KO). (B) S2KO cells were complemented by lentivirus transduction with SMYD2-WT (S2KO-S2WT) or GFP (S2KO-GFP) to generate polyclonal cell lines. These cell lines were compared by western blot with cells derived from those, which were additionally transformed with oncogenic H-RAS G12V (S2KO-HRAS-GFP, S2KO-HRAS-S2WT).

5 Discussion

5.1 SMYD2 and histone methylation

Many of the initially characterized mammalian protein methyltransferases were shown to methylate histones. This was the reason, why this protein class was first described as histone methyltransferases (HMTs) rather than protein methyltransferases (PMTs). Prominent representatives of the SET domain containing protein lysine methyltransferases (PKMTs) with known histone methyltransferase activity are EZH2 (H3 K27 methylation) [16] and members of the MLL family like MLL3/MLL4 (H3 K4 methylation) [52]. However, more recent data have also suggested that many of the PKMTs of the human genome are also able to methylate non-histone proteins [54]. Similar to other PKMTs, SMYD2 was initially reported to be a histone methyltransferase of H3 K4 and K36. However, these data were solely based on biochemical methylation assays using recombinant histones or reconstituted nucleosomes [1;13]. In the present study, global changes of these marks were not detected in cells after siRNA mediated downregulation of SMYD2 protein (**Fig. 4.9**) or after inhibition of SMYD2 enzyme activity using BAY-598 (data not shown). These observations are in line with analyses of global levels of all methylation states of H3 K4 and K36 in heart tissue from cardiomyocyte specific SMYD2 knockout mice [27], where the authors also did not observe significant changes in global histone methylation compared to control mice. Furthermore, another study in KYSE150 cells (SMYD2 amplified) did not detect significant differences in a whole panel of histone modifications using quantitative mass spectrometry after inhibition of SMYD2 activity with a small molecule inhibitor [93]. Although one cannot rule out the possibility that some of these marks may be affected locally on individual histones, so far cellular data are missing that would demonstrate a methylation activity of SMYD2 on histones. A technical obstacle to answer the question if SMYD2 may alter focal histone methylation is how to distinguish direct from indirect effects. For example, experimental interventions by either knockdown, knockout or small molecule inhibition to reduce SMYD2 activity may lead to changes in focal histone methylation under certain conditions. However, these changes could occur due to changes in gene expression patterns as a response of SMYD2 inhibition or directly due to interfering with SMYD2 activity on these histones. To better clarify this, chromatin immunoprecipitation experiments would be needed to map SMYD2 occupancy on the chromatin and align this with the occurrence of individual histone marks. Preferably, the ability to capture chromatin by immunoprecipitation of SMYD2 should be compared to knockout samples or after SMYD2 inhibition to demonstrate that obtained peaks are SMYD2 specific.

5.2 SMYD2 subcellular localization

Connected to the open question about in vivo histone methylation activity, controversial reports on the subcellular localization of SMYD2 exist. Whereas some reports observed nuclear localization, which would support potential histone methyltransferase activity, other work rather points to a cytoplasmic localization. In the present study, a clear cytoplasmic localization of endogenous SMYD2 was observed in NIH3T3 mouse fibroblasts in immunofluorescence images (**Fig. 4.46**). The signal was completely absent in NIH3T3 SMYD2 knockout cells, indicating the specificity of the antibody. Clear cytoplasmic localization of endogenous SMYD2 was also observed in KYSE150 cells (**Fig. S7.52**), and in HeLa and MDA-MB231 cells on western blot after subcellular fractionation (**Fig.**

S7.53) as well as predominantly cytoplasmic localization of ectopic SMYD2 in KYSE150 cells in immunofluorescence images (**Fig. 4.19**). Donlin et al. tested the localization of several GFP fused PMTs and reported predominantly cytoplasmic localization of ectopic SMYD2-GFP fusion proteins in HEK293 cells[30]. After quantification of GFP signals from immunofluorescence images, they measured over 80% of the signal in the cytoplasm. In comparison, Set6-GFP fusion protein, another methyltransferase was mainly localized to the nucleus and less than 10% of the signal derived from the cytoplasm, indicating the feasibility of this approach to determine subcellular distribution of PMTs. Also Voelkel et al. showed a clear cytoplasmic staining with a striated pattern of endogenous SMYD2 in cultured neonatal rat cardiomyocytes and a localization of SMYD2 in human myofibrills on either sides of the Z-disk in the I-band region[131]. In contrast, Diehl et al showed mainly nuclear localization of SMYD2 in cryosections of postnatal cardiomyocytes, which was completely absent in tissue from SMYD2-KO mice[27], demonstrating the specificity of detected signals. Also Cho et al. showed mainly nuclear SMYD2 localization in SBC5 cells, a lung adenocarcinoma cell line[19]. It is unclear so far, how to explain such different observations. The differences might be explained by different cell types tested or different conditions and might suggest a potential shuffling of SMYD2 between the nucleus and the cytoplasm. However, clear evidence is missing that would support an active shuffling between these compartments. Donlin et al. also noted, using in silico prediction algorithms that they did not identify nuclear localization signals (NLS) in the SMYD2 sequence but potential nuclear export signals (NES)[30]. Localization between different compartments could also be accomplished by the formation of different protein complexes, which distribute differentially in the cell. Clearly, more work is needed to clarify this question and although potential protein protein interaction partners for SMYD2 have been reported, it is unclear in what kind of complexes SMYD2 is functional.

5.3 SMYD2 in transcriptional regulation

Protein methyltransferases have been shown to regulate gene transcription. For instance, the methyltransferase EZH2 as part of the polycomb repressive complex 2 (PRC2) mediates H3 K27 mono-, di-, and trimethylation at gene promotors, which is commonly associated with gene silencing. In contrast, H3 K4 trimethylation, deposited by MLL containing COMPASS complexes[87] at gene promotors, is normally associated with gene activation, showing that histone methylation can promote both transcriptional repression or activation depending on the specific residue. More recent data have also pointed towards transcriptional regulation by lysine methylation beyond histone methylation by influencing transcription factor activity[45] or more indirect even by modulating cellular signaling pathways, as is the case for SMYD3[86]. Taken together, in one or the other way, protein methylation is often involved in transcriptional regulation. With that respect, also SMYD2 has been reported in the literature to regulate transcription, although in opposite ways and with different modes of action. For example, Brown et al. observed inhibition of gene expression in a reporter gene assay upon ectopic SMYD2 overexpression[13]. This was unexpected as in the same study SMYD2 was also observed to in vitro methylate H3 K36, a histone mark normally associated with transcriptional elongation[132]. The authors explained their observation by a potential association of SMYD2 with the repressive histone deacetylase complex Sin3A, but data on endogenous systems were missing. Huang et al. described an inhibition of p53 target gene transcription by SMYD2 after DNA damage. This was explained by p53 methylation at K370, which prevented p53 from efficiently binding to chromatin[55]. In contrast, overexpression of SMYD2 in HEK293 cells resulted predominantly in upregulation

of genes[1]. In this study, the authors detected ectopic SMYD2 occupancy in the promotor regions of some of the upregulated genes and speculated, based on biochemical methylation data, that SMYD2 mediated H3 K4 methylation and subsequent gene activation[1]. This is in contradiction to the predominant cytoplasmic localization of SMYD2 in HEK293 cells[30], but may suggest that a subpool of SMYD2, at least under ectopic overexpression, acts as nuclear H3 K4 methyltransferase and regulates gene activation.
In view of these observations and in the context of the two above discussed questions about SMYD2 histone methyltransferase activity and subcellular localization, I was interested, if and how SMYD2 might regulate gene transcription. To investigate on that question, 5 polyclonal cell lines from the MDA-MB231 breast cancer cell line model were generated, which stably expressed 5 different short hairpin oligonucleotids that efficiently mediated SMYD2 knockdown. The aim was to compare global gene expression in those cell lines with an empty vector control cell line to identify potential target genes regulated by SMYD2. However, despite reasonable expression of SMYD2 in the parental cell line, efficient knockdown did not reveal significant phenotypic changes. This was reflected by only moderate changes in global gene expression. More importantly, comparison of the individual SMYD2 knockdown cell lines revealed no commonly regulated genes besides SMYD2 itself, suggesting that changes in gene expression were rather associated with off-target effects of individual shRNAs. It has been appreciated that intervention by RNAi mediated knockdown can lead to off-target effects that can compromise the interpretation of observed phenotypes or gene expression changes. Indeed, subsequently during the project, in a similar experiment, parental MDA-MB231 cells were treated for 3 days with 1 µM BAY-598 (20x over cellular IC50 of substrate methylation) or DMSO and global gene expression analysis revealed no changes in gene expression at all (data not shown). These findings are similar to gene expression data from cardiomyocytes of SMYD2 knockout versus control mice, where only minimal changes were observed[27]. Taken together, the observations in the current study that SMYD2 was mainly localized in the cytoplasm, global histone methylation was unaffected, as was proliferation of knockdown cell lines, and no significant changes in gene expression were observed, suggest that SMYD2 is not a direct transcriptional regulator on chromatin at least under the tested conditions. This is in contrast to some of the previous reports and might be explained by different cell systems and different conditions tested. Perhaps, SMYD2 might have direct or indirect effects on transcription upon specific stress conditions or stimuli either by relocalization to the nucleus or by modulating signaling pathways that become activated by the stimuli.

5.4 SMYD2 dependent substrate methylation

Soon after the first characterization of SMYD2 as a histone methyltransferase, it was demonstrated that SMYD2 possesses additional methylation activity towards the tumor suppressor p53 - a finding that was also detected in cells[55]. This mono-methylation of p53 at lysine residue K370 was suggested to inhibit the transactivation activity of p53 and thereby inhibiting p53 tumor suppressor activity. To potentially use this methylation activity of SMYD2 for the development of a cellular mechanistic assay, a customized antibody (SY46) was generated to detect p53 monomethylation at lysine residue K370. Indeed, cross-purified SY46 antibody was able to specifically detect methylated p53 and this mark could also be decreased in cells upon SMYD2 inhibition. However, obtained signals from endogenous proteins were too weak for the development of a quantitative cellular mechanistic assay. Therefore, ectopic SMYD2 overexpression was intended to increase the signal. Interestingly however, ectopic SMYD2 did not increase detectable p53

methylation but revealed additional protein methylation. A reason for the still low levels of p53 methylation, despite high total p53 protein levels, could be that ectopic SMYD2 did not increase the number of physical contacts to p53 protein due to different localizations within the cell and that only a small subset of total p53 was actually in close proximity to SMYD2. Indeed, subcellular fractionation revealed nuclear and cytoplasmic distribution of p53 whereas SMYD2 was mainly cytoplasmic (**Fig. S7.52**). Nonetheless, the strong reactivity of SY46 antibody upon ectopic SMYD2 overexpression pointed towards additional SMYD2 substrates with epitopes similar to p53 at K370 site. In fact, other groups have reported on additional substrates besides p53 and histones like RB1[105], HSP90[30], PARP1[98], PTEN[91], and ERα[141]. However, none of the reported substrates were able to explain the observed methylation pattern at high molecular weight in KYSE150, MDA-MB231 and HeLa cells upon SMYD2 overexpression. Although the substrate(s) behind the observed methylation signal initially remained elusive, it served as a convenient opportunity to develop a robust quantitative assay to evaluate the cellular potency of synthezised small molecular SMYD2 inhibitors. This, among other project activities, led to the selection of the very potent and selective SMYD2 inhibitor probe BAY-598, which is now available to the research community via the Structural Genomics Consortium (SGC, `www.thesgc.org/chemical-probes/BAY-598`) and will hopefully allow scientists to enhance our understanding of the mysterious biological function of SMYD2.

In addition to successfully having supported the development of a SMYD2 probe inhibitor, I aimed at identifying novel SMYD2 substrates by measuring proteome wide lysine methylation in MDA-MB231 cells ectopically overexpressing SMYD2 with the hypothesis that ectopic SMYD2 would have a significant impact on the overall lysine methylation. In total 188 and 695 sites of lysine mono-methylation were detected with SY46 and a pan-Kme1 antibody, respectively. This demonstrated the general feasibility of this approach to identify novel cellular lysine methylation and clearly shows how less lysine methylation outside of histones is recognized and understood. Surprisingly, from all previous reported SMYD2 substrates only HSP90 K615 methylation was detected. This could have several reasons: First, some of the described substrates were probably not sufficiently expressed at the protein level in the MDA-MB231 cell line model. Indeed, the described substrate estrogen receptor alpha (ERα)[60;141] is not expressed in the ER negative MDA-MB231 cell line[63] and would therefore not be expected to be detected. Another reason could be that the antibody was not able to capture some of the previous described methylation site due to different physicochemical properties. For instance, the amino acid composition around the described methylation site of PTEN at K313 significantly differs from that of p53 at K370. Third, a neighboring modification like acetylation or phosphorylation could have prevented detection of the methylation site by the antibody. For example, it was observed that phosphorylation at serine residue S28 can prevent detection of trimethylation at K27 by an H3 K27me3 specific antibody using histone peptide arrays with combinations of different modifications (inhouse observations). Last but not least, certain methylation sites may be specifically regulated and did not occur under the tested conditions although protein abundance of both enzyme and target protein were high. For instance, gene expression data suggested high PARP1 abundance in MDA-MB231 cells. However, no methylation was detected, perhaps, because PARP1 is involved in DNA single-stranded breaks repair[68] and only under those conditions methylation might be relevant.

Nonetheless, with the help of stringent selection criteria, I was able to identify and validate

AHNAK as a novel SMYD2 substrate that had the most dominant impact on the observed methylation signals detected by SY46 antibody. AHNAK methylation was most clearly detected upon ectopic SMYD2 overexpression but was also measurable at endogenous levels of SMYD2 both in human cancer cell lines as well as in mouse cell lines and muscle tissue from mice. This suggests that AHNAK methylation is conserved in different species. Moreover, AHNAK methylation provides an unprecedented example of multi-methylation due to its repetitive domain structure. In the present proteome study using MDA-MB231 breast cancer cells ectopically overexpressing SMYD2, 79 unique mono-methyl lysine sites were detected with SY46 and a pan-Kme1 antibody together. It is unclear though, if all methylation sites detected within AHNAK can be assigned to SMYD2 activity, but a sequence motif derived from all AHNAK methylation sites showed a clear enrichment for a LKGPK motif. This suggests that at least lysines within the LKGPK or a similar motif are likely target sites for SMYD2 methylation. The typical validation procedure, in which single lysine residues are mutated to arginine to precisely map the methylation sites, was unfeasible in this case given the high number of potential target sites and the large size of the protein. However, in a biochemical methylation assay using recombinant AHNAK-CRUs and recombinant SMYD2 enzyme it could be demonstrated that SMYD2 can directly methylate recombinant AHNAK protein fragments. Subsequent mass spectrometry analyses from AHNAK-CRUs excised from protein gels detected peptides comprising the methylated LK(me1)GPK motif and to a lesser extend LKGPK(me1) only in the samples that had been incubated with SMYD2. This not only demonstrated that SMYD2 is able to directly methylate AHNAK-CRUs but also further supported the idea that LKGPK is a preferred motif for SMYD2 methylation. Surprisingly though, a biochemical SPA methylation assay using a peptide derived from AHNAK that comprised the LKGPK motif failed to reveal measurable methylation in contrast to p53 peptide, which was efficiently methylated under these conditions. One reason for such a negative result could be the differences in the biotinylation sites of the peptides, as AHNAK derived peptide was biotinylated at the C-terminus whereas p53 peptide was tagged with Biotin at the N-terminus. Interestingly, Wu et al. compared the methyltransferase activity of SMYD2 in biochemical assays using full-length p53, histones and peptides thereof[137]. They found that actually full-length histone H2B and histone H4 were better substrates than p53 protein, which itself was a better substrate than histone H3. However, when comparing peptides derived from these proteins, p53 peptide comprising K370 showed significant better methylation rates than any other peptides from histones. Using multistate computational protein design algorithms, Lanouette et al. tried to predict novel SMYD2 substrates using the previously described p53 peptide as starting point for their model[74]. They complemented this approach with a biochemical methylation screen using permutated peptide substrates and measured the methylation efficiency. With both approaches they hypothesized a recognition motif for SMYD2 methylation of [LFM]-K(me1)-[AFYMSHRK]-[LYK]. Although there are similarities between the recognition motifs from AHNAK methylation and from Lanouette et al., especially at the -1 position, glycine and proline at +1 and +2 position, respectively were excluded by Lanouette et al. but were the dominant amino acids in the sequence motif of AHNAK methylation. This shows that there is obviously a discrepancy in the methylation preferences between unstructured oligopeptides and structured polypeptides or proteins. On one hand, it means that the potential spectrum of SMYD2 substrate methylation sites might be broader than originally predicted, on the other hand the presence of certain recognition motifs and the ability for in vitro methylation of peptides does not allow a 1:1 inference to in vivo methylation. To be mentioned, none of the predicted substrates from Lanouette et al. were either detected in the present proteome study or in the study

from Olsen et al.[96] (discussed below). Reasons are diverse and can be due to technical limitations as already discussed above, but will also be due to the cellular context and spatial proximity of enzyme and substrate within the cell.

5.5 Potential implications of AHNAK methylation

Having identified AHNAK as a substrate of SMYD2 that can be multi-methylated on many lysine residues with LKGPK as the preferred methylation motif, the perhaps most interesting question remains: What is the biological relevance of AHNAK methylation?
AHNAK, which means giant in Hebrew, is indeed a "giant" protein of about 630 kDa (UniProt ID Q09666). It consists of an N-terminal PDZ domain, a large central domain of 4300 aa, which itself is made of approximately 38 central repeated units of 68-128 aa in length, and a C-terminal domain with potential nuclear localization motifs. The PDZ domain is a protein-protein binding module implicated in binding C-terminal regions of membrane receptor proteins (e.g. G protein coupled receptors (GPCRs)) and connecting these to the cytoskeleton[76].
AHNAK lacks enzymatic activity and its concrete function remains elusive. It most likely acts as a scaffold protein to form multi-protein complexes. Most reports implicated regulation in cell signaling and calcium channels[24], however, more recent publications have suggested also implications in epithelial to mesenchymal transition (EMT) and tumor metastasis[114;125]. It was first described by Shtivelmann et al. to be an unusually large protein with a polyionic rod like shape of 1.2 µm and to be differentially expressed in neuroblastomas compared to healthy nerve tissue[119]. At about the same time, Hashimoto et al. discovered a large protein of desmosomal plaques, which was therefore named Desmoyokin, and turned out to be identical to AHNAK[50]. Similar to SMYD2, AHNAK knockout mice are viable, fertile and show no obvious phenotype[71]. Later, Haase et al. found a large phosphoprotein, called phosphoprotein 700 (pp700), that was associated with the beta 2 subunit of cardiac L-type calcium channels in porcine and rat cardiomyocytes, to be the ortholog to human AHNAK[43]. Phosphorylation of AHNAK occurred in a cAMP dependent manner by protein kinase A (PKA) after beta-adrenergic receptor stimulation. It was suggested that under basal conditions AHNAK inhibits calcium channel activation, but phosphorylation of AHNAK releases the beta-2 subunit of the calcium channel, which allowed calcium influx and contraction. PKA is a serine/threonine kinase. In fact, many serine residues lie within the CRUs of AHNAK in the vicinity of detected methylated lysine residues. Therefore, one might speculate that these serine residues might be phosphorylated by PKA, depending on the presence or absence of nearby lysine methylation, or, like in the case of MAP3K2 methylation by SMYD3[86], that phosphorylation is maintained by nearby methylation through a mechanism that prevents phosphatase recruitment.
AHNAK was also found to bind and activate phospholipase C (PLC) in the presence of arachidonic acid[111] as well as protein kinase C (PKC)[77;78]. Interestingly, binding and activation of both PLC and PKC was mediated by the central repeated units of AHNAK. Activation of PLC was also connected to breast cancer metastasis[108]. If methylation of AHNAK influences PLC or PKC binding or activation might be worth investigating.
Immunofluorescence images of ectopic SMYD2 and SY46 detected methylation in KYSE150 cells revealed mainly cytoplasmic staining in interphase cells. However, both SMYD2 as well as methylation signals revealed strong staining of the spindle machinery in mitotic cells. Localization of SMYD2 to spindle fibers has not yet been described. Interestingly, mitotic spindle localization was detected by several monoclonal antibodies against the PEVK region of titin (the largest protein of the human genome and specific to muscle

cells) in non-muscle derived cells[135]. In the same study, the authors demonstrated that these antibodies cross-reacted with AHNAK protein, suggesting that SMYD2 together with methylated AHNAK are localized to and might also regulate mitotic spindles. Definitively, more work is needed to reveal a clear function of SMYD2 and AHNAK methylation.
However, answering this question in the future will be challenging as knockout mice of both SMYD2 and AHNAK did not show obvious phenotypes[27;72;103]. If a protein as a whole is dispensable for normal development a phenotype of a posttranslational modification on that protein might be even more subtle. It also raises the question, why are both SMYD2 enzyme as well as methylation motifs on AHNAK conserved among species, if they are dispensable et least under laboratory conditions? One explanation could be that there is functional redundancy which is able to compensate loss of function of both proteins. However, redundancy is unlikely to be accomplished by paralogs as there was no indication for upregulation of other SMYD family members upon SMYD2 inhibition and AHNAK and AHNAK2 seemed to be reversely regulated at least during C2C12 differentiation. My personal view on that question is that both may be relevant under certain stress conditions so far not tested or would give an evolutionary advantage in the long run in a real world environment. More work is definitively needed to uncover this mystery.

5.6 Comparison with another proteome study of SMYD2 in KYSE150

Given the high number of detected methylation sites in the current proteome study and the obviously brougth substrate specificity of SMYD2, it is very likely that SMYD2 activity was responsible for additional methylation sites of other proteins, although this might have been overshadowed by AHNAK methylation. However, to clearly identify those, more comprehensive quantitative methods would have been required. Interestingly, very recently, Olsen et al. have reported on results of another proteome study in the KYSE150 cell line model[96]. In that study, they indeed used a quantitative SILAC (stable isotope labeling of amino acids[97]) based proteomic approach and compared global lysine methylation of SMYD2 knockdown cells with the corresponding parental cell line and in addition, they also compared global lysine methylation in ectopic SMYD2 overexpressing KYSE150 cells before and after treatment with the SMYD2 inhibitor probe LLY-507[93]. By that, they were able to quantitatively measure relative changes in protein methylation between treatment groups. They identified 35 methylation sites that were significantly reduced upon SMYD2 depletion/inhibition. Strikingly, very similar to my own results, their list of 35 regulated methylation sites included HSP90 methylation at K615, but none of the other reported SMYD2 substrates. What is more, out of the 35 methylation sites identified by Olsen et al. 17 sites were also detected in the present proteome study (for details see supplement). Most importantly, among them also several sites of AHNAK and AHNAK2. This really indicates that SMYD2 contains a larger set of substrates that are common to cells of different tissue origin and that AHNAK and AHNAK2 are among these commonly chaired substrates.
It is now tempting to speculate if there is a functional connection between the different substrates. The 17 sites are within proteins like MAP2K4[80], a family member of the mitogen activated kinases that is able to activate the stress activated JNK1, JNK2, and p38 kinases.
FIGNL1, a gene associated with DNA damage response where it plays a role in DNA double-strand break repair via homologous recombination[140].
HN1 also contained a methylation site that was detected in MDA-MB231 and for which

downregulation of HN1 methylation was reported in KYSE150 cells upon SMYD2 inhibition. This protein has been implicated in regulation of regeneration and differentiation and was connected to the p38 kinase pathway.
Another protein PHACTR2 has been described to regulate the activity of protein phosphatase 1 (PP1) and actin[4]. PP1 is a ubiquitously expressed protein phosphatase that is modulated in its dephosphorylation activity towards specific substrates by association with different regulatory subunits like PHACTR2. Methylation of PHACTR2 could potentially influence its association with PP1. However, this would need further investigation.
Another potential SMYD2 target detected in both cell lines is PRKDC. This gene encodes the catalytic subunit of the DNA-dependent protein kinase and is also involved in DNA double strand break repair and recombination.
Another prominent member of methylated proteins that was detected in KYSE150 and MDA-MB231, and which is presumably a substrate for SMYD2 methylation, is RICTOR, a subunit of the mTORC2 complex that is involved in phosphorylation of AKT at S473 and in cytoskeleton regulation through its stimulation of F-actin stress fibers[57;112]. These studies have suggested that mTORC2 controls the actin cytoskeleton by promoting protein kinase C alpha (PKCα) phosphorylation, phosphorylation of paxillin and its relocalization to focal adhesions, and the GTP loading of RhoA and Rac1. The molecular mechanism by which mTORC2 regulates these processes has not been determined, but interestingly, as mentioned earlier, AHNAK was suggested to activate PKC. To mention in this context, AHNAK was described to be localized to costameres in muscle cells[85], which can be viewed as specialized focal adhesions and share many common proteins.

5.7 Potential functions of SMYD2 mediated lysine methylation

Several functions have been assigned to lysine methylation mostly driven by observations on histone methylation, but these may also be transferred to non-histone methylation. The fact, that in endogenous systems only small fractions of total substrates seem to be methylated, suggests that protein methylation similar to histone methylation allows for differentiated functions of sub-pools of proteins.
Lysine residues are targets of a number of posttranslational modifications besides methylation such as acetylation, ubiquitination or sumoylation among others[82]. Importantly, these different modifications are mutually exclusive and possess different properties. That means different modifications may be competing against each other on certain lysine residues and promote different outcomes. For example, lysine methylation may prevent polyubiquitination, which otherwise would lead to protein degradation through the ubiquitin-proteasome pathway. However, there was no evidence so far that this could be the case for AHNAK, because no other lysine modifications for this protein have been reported to date and protein stability seemed not to be influenced in the present experiments after changes of lysine methylation.
Another function of lysine methylation is the modulation of protein-protein interactions. Lysine methylation does not change the charge of the protein in contrast to phosphorylation or acetylation and leads only to a small change in size (14 Da). However, it alters the hydration energies and hydrogen bonding potential of the lysine side chains[25]. There are a number of so called "reader" domains that have been identified, which allow the recruitment of the corresponding proteins to the site of lysine methylation and thereby enabling the assembly of functional protein complexes. Domains with reported methyl binding properties are TUDOR[84], PHD[14], MBT[11], CHROMO[33], PWWP[100] and WD40 repeat[126] domains. Therefore, one could envision that, in analogy to chromatin, AHNAK

serves as a scaffold for the recruitment and assembly of specific signalosomes, which are regulated by diverse combinations of posttranslational modifications on the protein, among them SMYD2 mediated lysine methylation. It is unclear though, if the above mentioned domains for lysine methylation recognition are also relevant in the cytoplasm or if other domains may be able to do so.
In addition, another possibility is that lysine methylation prevents protein interaction. For instance MAP3K2 protein kinase has been reported to be methylated by SMYD3, a close homolog of SMYD2. This methylation prevented interaction with the phosphatase PP2A and thereby maintaining MAP3K2 phosphorylation and activation[86]. AHNAK has been described to be a phospho-protein[118] and several serine residues lie in the vicinity to the methylated lysine residues. Therefore, one could imagine that there exists a cross talk between both modifications. Unfortunately, I was not able to detect AHNAK phosphorylation with a pan-Ser/Thr antibody (data not shown) and therefore could not pursue on that hypothesis. Another way to determine if AHNAK methylation influences protein-protein interactions would be a GST-pulldown with recombinant AHNAK or subunits thereof, that are either methylated or unmethylated and test, whether certain proteins are differentially bound to both forms of AHNAK. This approach proved to be successful for identifying the relevance of MAP3K2 methylation. However, such approach would require quantitative SILAC based mass spectrometry analyses and efficient methylation of recombinant AHNAK-CRUs in order to obtain a clear differentiation.

5.8 Functional characterization of SMYD2 inhibition

In the last part of my thesis I functionally characterized the effects of SMYD2 inhibition or depletion in selected cell models. The first model system was chosen to reveal the relevance of SMYD2 inhibition in normal physiology. For this, the C2C12 cell line model was selected because of the observation of high SMYD2 protein expression in skeletal muscle tissue. C2C12 are mouse myoblasts that can be easily differentiated into myotubes and form muscle fiber like structures that even start to contract after several days of differentiation. Analysis of gene expression revealed a successive downregulation of SMYD2 during differentiation. This was unexpected, as high levels of SMYD2 in muscle tissue of adult mice suggested also high expression in differentiated C2C12 myotubes. However, it is difficult to compare both systems and perhaps SMYD2 expression in muscle precursors during development or undifferentiated satellite cells within adult muscle tissue is even higher than in differentiated muscle fibers. Despite significant expression of SMYD2, permanent inhibition of SMYD2 activity with BAY-598 treatment had no measurable impact on myoblasts and differentiation. This is actually in line with the fact that SMYD2 knockout mice are viable and show no developmental defects even in tissues of high SMYD2 expression like heart and skeletal muscle. The lack of phenotype may be explained by compensatory mechanisms or redundancy. However, analyses of transcript and protein expression of the other SMYD family members gave no hint that compensation might be achieved by close paralogs. Also literature does not point sofar to overlapping substrates of different SMYD proteins, although a systematic approach for identifying substrates for the other SMYD family members are still missing and therefore shared substrates cannot be ruled out. Stable knockdown of SMYD2 in MDA-MB231 cells did also not reveal consistent upregulation of other genes in different knockdown cell lines in a global gene expression study.

5.8.1 SMYD2 activity does not impact H-Ras transformation of NIH3T3 cells

The second model system was intended to reveal more insights about SMYD2 function during tumorigenesis based on a report that observed SMYD2 upregulation in a mouse model of oncogenic RAS driven ductal adenocarcinoma formation in pancreas[103]. Strikingly, conditional knockout of SMYD2 in pancreas had no impact on normal pancreas physiology but significantly compromised oncogenic RAS driven tumor formation in this organ. This suggested that SMYD2 might be an important downstream effector of oncogenic RAS signaling. To recapitulate this idea in a simple cellular in vitro system, NIH3T3 mouse fibroblasts were employed that are highly succeptable for oncogene transformation, especially oncogenic RAS. For that purpose I not only generated NIH3T3 SMYD2-KO cells, in analogy above mentioned publication but also tested in parallel the clinically more relevant inhibition of SMYD2 via BAY-598. However, in both cases permanent inhibition of SMYD2 actitiy with BAY-598 or via CrispR/Cas9 knockout of SMYD2 did not compromise H-RAS G12V transformation of NIH3T3. Also parental NIH3T3 cells did not reveal protein upregulation of SMYD2 after RAS transformation. This indicated that SMYD2 is probably not a direct downstream effector of RAS signaling. The observed effects in pancreas in vivo points towards more complex regulation that cannot be easily reconstructed in 2D cell culture. Therefore, I suspect that upregulation of SMYD2 in pancreatic lesions was not a direct effect of RAS signaling but rather an indirect consequence of dramatic changes in the microenvironment of cancerous lesions, like inflammation, immune cell infiltration, hypoxia and other stresses. The authors also noted decreased inflammatory response in SMYD2-KO mice during tumor formation, especially lower serum levels of IL-6, a proinflammatory cytokine. Therefore, it might be interesting to further investigate on the role of SMYD2 during inflammation, a process that is not activated in normal development, which might explain a lack of phenotype.

5.9 Conclusion

The current work constitutes a significant contribution to the development of the specific small molecule SMYD2 inhibitor probe BAY-598 by establishing a quantitative mechanistic assay that was able to measure the methyltransferase activity of this enzyme in cells. Furthermore, I was able to identify and validate AHNAK as a novel SMYD2 substrate, which showed unprecedented multi-methylation along the protein. Although the biological relevance of SMYD2 and AHNAK methylation or other SMYD2 substrate methylation is still mysterious, initial tests of SMYD2 inhibition in two cell culture models showed the feasibility of inhibitor probes to test functional hypotheses. Therefore, with the help of BAY-598 it will be easy and convenient to test more sophisticated model systems, which are obviously needed to uncover SMYD2 function. Coming back to the initial idea that SMYD2 might be a promising target in cancer therapy, when looking at the present data in conjunction with findings from Eggert et al.[32] it appears rather unlikely to employ SMYD2 inhibition in a mono-therapy strategy in cancer patients. However, future research may identify rational combination therapies.

6 References

[1] M. Abu-Farha, J. P. Lambert, A. S. Al-Madhoun, F. Elisma, I. S. Skerjanc, and D. Figeys. The tale of two domains: proteomics and genomics analysis of SMYD2, a new histone methyltransferase. *Mol. Cell Proteomics*, 7(3):560–572, Mar 2008. [DOI:10.1074/mcp.M700271-MCP200] [PubMed:18065756].

[2] M. Abu-Farha, S. Lanouette, F. Elisma, V. Tremblay, J. Butson, D. Figeys, and J. F. Couture. Proteomic analyses of the SMYD family interactomes identify HSP90 as a novel target for SMYD2. *J Mol Cell Biol*, 3(5):301–308, Oct 2011. [DOI:10.1093/jmcb/mjr025] [PubMed:22028380].

[3] US Food & Drug Administration. *How Drugs are Developed and Approved*, 2016. `http://www.fda.gov/Drugs/DevelopmentApprovalProcess/HowDrugsareDevelopedandApproved/default.htm` [Updated: August 18, 2015].

[4] P. B. Allen, A. T. Greenfield, P. Svenningsson, D. C. Haspeslagh, and P. Greengard. Phactrs 1-4: A family of protein phosphatase 1 and actin regulatory proteins. *Proc. Natl. Acad. Sci. U.S.A.*, 101(18):7187–7192, May 2004. [PubMed Central:PMC406487] [DOI:10.1073/pnas.0401673101] [PubMed:15107502].

[5] R. P. AMBLER and M. W. REES. Epsilon-N-Methyl-lysine in bacterial flagellar protein. *Nature*, 184:56–57, Jul 1959. [PubMed:13793118].

[6] C. H. Arrowsmith, C. Bountra, P. V. Fish, K. Lee, and M. Schapira. Epigenetic protein families: a new frontier for drug discovery. *Nat Rev Drug Discov*, 11(5):384–400, Apr 2012. [DOI:10.1038/nrd3674] [PubMed:22498752].

[7] D. E. Atkinson. *Cellular Energy Metabolism and its Regulation*, volume 3. Academic Press, 1977.

[8] A. J. Bannister, R. Schneider, and T. Kouzarides. Histone methylation: dynamic or static? *Cell*, 109(7):801–806, Jun 2002. [PubMed:12110177].

[9] T. A. Baudino. Targeted Cancer Therapy: The Next Generation of Cancer Treatment. *Curr Drug Discov Technol*, 12(1):3–20, 2015. [PubMed:26033233].

[10] M. T. Bedford and S. G. Clarke. Protein arginine methylation in mammals: who, what, and why. *Mol. Cell*, 33(1):1–13, Jan 2009. [PubMed Central:PMC3372459] [DOI:10.1016/j.molcel.2008.12.013] [PubMed:19150423].

[11] R. Bonasio, E. Lecona, and D. Reinberg. MBT domain proteins in development and disease. *Semin. Cell Dev. Biol.*, 21(2):221–230, Apr 2010. [PubMed Central:PMC3772645] [DOI:10.1016/j.semcdb.2009.09.010] [PubMed:19778625].

[12] P. A. Boriack-Sjodin and K. K. Swinger. Protein Methyltransferases: A Distinct, Diverse, and Dynamic Family of Enzymes. *Biochemistry*, 55(11):1557–1569, Mar 2016. [DOI:10.1021/acs.biochem.5b01129] [PubMed:26652298].

[13] M. A. Brown, R. J. Sims, P. D. Gottlieb, and P. W. Tucker. Identification and characterization of Smyd2: a split SET/MYND domain-containing histone H3 lysine 36-specific methyltransferase that interacts with the Sin3 histone deacetylase complex. *Mol. Cancer*, 5:26, Jun 2006. [PubMed Central:PMC1524980] [DOI:10.1186/1476-4598-5-26] [PubMed:16805913].

[14] T. R. Burglin and M. Affolter. Homeodomain proteins: an update. *Chromosoma*, 125(3):497–521, Jun 2016. [PubMed Central:PMC4901127] [DOI:10.1007/s00412-015-0543-8] [PubMed:26464018].

[15] E. Calpena, F. Palau, C. Espinos, and M. I. Galindo. Evolutionary History of the Smyd Gene Family in Metazoans: A Framework to Identify the Orthologs of Human Smyd Genes in Drosophila and Other Animal Species. *PLoS ONE*, 10(7):e0134106, 2015. [PubMed Central:PMC4521844] [DOI:10.1371/journal.pone.0134106] [PubMed:26230726].

[16] R. Cao, L. Wang, H. Wang, L. Xia, H. Erdjument-Bromage, P. Tempst, R. S. Jones, and Y. Zhang. Role of histone H3 lysine 27 methylation in Polycomb-group silencing. *Science*, 298(5595):1039–1043, Nov 2002. [DOI:10.1126/science.1076997] [PubMed:12351676].

[17] L. B. Chen, J. Y. Xu, Z. Yang, and G. B. Wang. Silencing SMYD3 in hepatoma demethylates RIZI promoter induces apoptosis and inhibits cell proliferation and migration. *World J. Gastroenterol.*, 13(43):5718–5724, Nov 2007. [PubMed Central:PMC4171257] [PubMed:17963297].

[18] X. Cheng, R. E. Collins, and X. Zhang. Structural and sequence motifs of protein (histone) methylation enzymes. *Annu Rev Biophys Biomol Struct*, 34:267–294, 2005. [PubMed Central:PMC2733851] [DOI:10.1146/annurev.biophys.34.040204.144452] [PubMed:15869391].

[19] H. S. Cho, S. Hayami, G. Toyokawa, K. Maejima, Y. Yamane, T. Suzuki, N. Dohmae, M. Kogure, D. Kang, D. E. Neal, B. A. Ponder, H. Yamaue, Y. Nakamura, and R. Hamamoto. RB1 methylation by SMYD2 enhances cell cycle progression through an increase of RB1 phosphorylation. *Neoplasia*, 14(6):476–486, Jun 2012. [PubMed Central:PMC3394190] [PubMed:22787429].

[20] S. Clarke. Protein methylation. *Curr. Opin. Cell Biol.*, 5(6):977–983, Dec 1993. [PubMed:8129951].

[21] J. F. Couture, E. Collazo, J. S. Brunzelle, and R. C. Trievel. Structural and functional analysis of SET8, a histone H4 Lys-20 methyltransferase. *Genes Dev.*, 19(12):1455–1465, Jun 2005. [PubMed Central:PMC1151662] [DOI:10.1101/gad.1318405] [PubMed:15933070].

[22] G. E. Crooks, G. Hon, J. M. Chandonia, and S. E. Brenner. WebLogo: a sequence logo generator. *Genome Res.*, 14(6):1188–1190, Jun 2004. [PubMed Central:PMC419797] [DOI:10.1101/gr.849004] [PubMed:15173120].

[23] C. Curtis, S. P. Shah, S. F. Chin, G. Turashvili, O. M. Rueda, M. J. Dunning, D. Speed, A. G. Lynch, S. Samarajiwa, Y. Yuan, S. Graf, G. Ha, G. Haffari, A. Bashashati, R. Russell, S. McKinney, A. Langerod, A. Green, E. Provenzano, G. Wishart, S. Pinder, P. Watson, F. Markowetz, L. Murphy, I. Ellis, A. Purushotham, A. L. Borresen-Dale, J. D. Brenton, S. Tavare, C. Caldas, S. Aparicio, C. Caldas, S. Aparicio, C. Curtis, S. P. Shah, C. Caldas, S. Aparicio, J. D. Brenton, I. Ellis, D. Huntsman, S. Pinder, A. Purushotham, L. Murphy, C. Caldas, S. Aparicio, C. Caldas, H. Bardwell, S. F. Chin, C. Curtis, Z. Ding, S. Graf, L. Jones, B. Liu, A. G. Lynch, I. Papatheodorou, S. J. Sammut, G. Wishart, S. Aparicio, S. Chia, K. Gelmon, D. Huntsman, S. McKinney, C. Speers, G. Turashvili, P. Watson, I. Ellis,

R. Blamey, A. Green, D. Macmillan, E. Rakha, A. Purushotham, C. Gillett, A. Grigoriadis, S. Pinder, E. de Rinaldis, A. Tutt, L. Murphy, M. Parisien, S. Troup, C. Caldas, S. F. Chin, D. Chan, C. Fielding, A. T. Maia, S. McGuire, M. Osborne, S. M. Sayalero, I. Spiteri, J. Hadfield, S. Aparicio, G. Turashvili, L. Bell, K. Chow, N. Gale, D. Huntsman, M. Kovalik, Y. Ng, L. Prentice, C. Caldas, S. Tavare, C. Curtis, M. J. Dunning, S. Graf, A. G. Lynch, O. M. Rueda, R. Russell, S. Samarajiwa, D. Speed, F. Markowetz, Y. Yuan, J. D. Brenton, S. Aparicio, S. P. Shah, A. Bashashati, G. Ha, G. Haffari, and S. McKinney. The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. *Nature*, 486(7403):346–352, Apr 2012. [PubMed Central:PMC3440846] [DOI:10.1038/nature10983] [PubMed:22522925].

[24] T. A. Davis, B. Loos, and A. M. Engelbrecht. AHNAK: the giant jack of all trades. *Cell. Signal.*, 26(12):2683–2693, Dec 2014. [DOI:10.1016/j.cellsig.2014.08.017] [PubMed:25172424].

[25] K. D. Daze and F. Hof. The cation-Ï€ interaction at protein-protein interaction interfaces: developing and learning from synthetic mimics of proteins that bind methylated lysines. *Acc. Chem. Res.*, 46(4):937–945, Apr 2013. [DOI:10.1021/ar300072g] [PubMed:22724379].

[26] P. A. Del Rizzo and R. C. Trievel. Substrate and product specificities of SET domain methyltransferases. *Epigenetics*, 6(9):1059–1067, Sep 2011. [PubMed Central:PMC3225744] [DOI:10.4161/epi.6.9.16069] [PubMed:21847010].

[27] F. Diehl, M. A. Brown, M. J. van Amerongen, T. Novoyatleva, A. Wietelmann, J. Harriss, F. Ferrazzi, T. Bottger, R. P. Harvey, P. W. Tucker, and F. B. Engel. Cardiac deletion of Smyd2 is dispensable for mouse heart development. *PLoS ONE*, 5(3):e9748, Mar 2010. [PubMed Central:PMC2840034] [DOI:10.1371/journal.pone.0009748] [PubMed:20305823].

[28] S. C. Dillon, X. Zhang, R. C. Trievel, and X. Cheng. The SET-domain protein superfamily: protein lysine methyltransferases. *Genome Biol.*, 6(8):227, 2005. [PubMed Central:PMC1273623] [DOI:10.1186/gb-2005-6-8-227] [PubMed:16086857].

[29] J. A. DiMasi, H. G. Grabowski, and R. W. Hansen. The cost of drug development. *N. Engl. J. Med.*, 372(20):1972, May 2015. [DOI:10.1056/NEJMc1504317] [PubMed:25970070].

[30] L. T. Donlin, C. Andresen, S. Just, E. Rudensky, C. T. Pappas, M. Kruger, E. Y. Jacobs, A. Unger, A. Zieseniss, M. W. Dobenecker, T. Voelkel, B. T. Chait, C. C. Gregorio, W. Rottbauer, A. Tarakhovsky, and W. A. Linke. Smyd2 controls cytoplasmic lysine methylation of Hsp90 and myofilament organization. *Genes Dev.*, 26(2):114–119, Jan 2012. [PubMed Central:PMC3273835] [DOI:10.1101/gad.177758.111] [PubMed:22241783].

[31] S. J. Du, X. Tan, and J. Zhang. SMYD proteins: key regulators in skeletal and cardiac muscle development and function. *Anat Rec (Hoboken)*, 297(9):1650–1662, Sep 2014. [DOI:10.1002/ar.22972] [PubMed:25125178].

[32] E. Eggert, R. C. Hillig, S. Koehr, D. Stockigt, J. Weiske, N. Barak, J. Mowat, T. Brumby, C. D. Christ, A. Ter Laak, T. Lang, A. E. Fernandez-Montalvan, V. Badock, H. Weinmann, I. V. Hartung, D. Barsyte-Lovejoy, M. Szewczyk, S. Kennedy, F. Li, M. Vedadi, P. J. Brown, V. Santhakumar, C. H. Arrowsmith, T. Stellfeld, and C. Stresemann. Discovery and Characterization of a

Highly Potent and Selective Aminopyrazoline-Based in Vivo Probe (BAY-598) for the Protein Lysine Methyltransferase SMYD2. *J. Med. Chem.*, 59(10):4578–4600, May 2016. [PubMed Central:PMC4917279] [DOI:10.1021/acs.jmedchem.5b01890] [PubMed:27075367].

[33] J. C. Eissenberg. Molecular biology of the chromo domain: an ancient chromatin module comes of age. *Gene*, 275(1):19–29, Sep 2001. [PubMed:11574148].

[34] N. L. Estrella, C. A. Desjardins, S. E. Nocco, A. L. Clark, Y. Maksimenko, and F. J. Naya. MEF2 transcription factors regulate distinct gene programs in mammalian skeletal muscle differentiation. *J. Biol. Chem.*, 290(2):1256–1268, Jan 2015. [PubMed Central:PMC4294490] [DOI:10.1074/jbc.M114.589838] [PubMed:25416778].

[35] Q. Feng, H. Wang, H. H. Ng, H. Erdjument-Bromage, P. Tempst, K. Struhl, and Y. Zhang. Methylation of H3-lysine 79 is mediated by a new family of HMTases without a SET domain. *Curr. Biol.*, 12(12):1052–1058, Jun 2002. [PubMed:12123582].

[36] A. D. Ferguson, N. A. Larsen, T. Howard, H. Pollard, I. Green, C. Grande, T. Cheung, R. Garcia-Arenas, S. Cowen, J. Wu, R. Godin, H. Chen, and N. Keen. Structural basis of substrate methylation and inhibition of SMYD2. *Structure*, 19(9):1262–1273, Sep 2011. [DOI:10.1016/j.str.2011.06.011] [PubMed:21782458].

[37] M. Fodinger, W. H. Horl, and G. Sunder-Plassmann. Molecular biology of 5,10-methylenetetrahydrofolate reductase. *J. Nephrol.*, 13(1):20–33, 2000. [PubMed:10720211].

[38] T. Fujii, S. Tsunesumi, H. Sagara, M. Munakata, Y. Hisaki, T. Sekiya, Y. Furukawa, K. Sakamoto, and S. Watanabe. Smyd5 plays pivotal roles in both primitive and definitive hematopoiesis during zebrafish embryogenesis. *Sci Rep*, 6:29157, Jul 2016. [PubMed Central:PMC4932602] [DOI:10.1038/srep29157] [PubMed:27377701].

[39] T. Fujii, S. Tsunesumi, K. Yamaguchi, S. Watanabe, and Y. Furukawa. Smyd3 is required for the development of cardiac and skeletal muscle in zebrafish. *PLoS ONE*, 6(8):e23491, 2011. [PubMed Central:PMC3160858] [DOI:10.1371/journal.pone.0023491] [PubMed:21887258].

[40] M. M. Ghahramani Seno, C. Trollet, T. Athanasopoulos, I. R. Graham, P. Hu, and G. Dickson. Transcriptomic analysis of dystrophin RNAi knockdown reveals a central role for dystrophin in muscle differentiation and contractile apparatus organization. *BMC Genomics*, 11:345, Jun 2010. [PubMed Central:PMC2890566] [DOI:10.1186/1471-2164-11-345] [PubMed:20515474].

[41] P. D. Gottlieb, S. A. Pierce, R. J. Sims, H. Yamagishi, E. K. Weihe, J. V. Harriss, S. D. Maika, W. A. Kuziel, H. L. King, E. N. Olson, O. Nakagawa, and D. Srivastava. Bop encodes a muscle-restricted protein containing MYND and SET domains and is essential for cardiac differentiation and morphogenesis. *Nat. Genet.*, 31(1):25–32, May 2002. [DOI:10.1038/ng866] [PubMed:11923873].

[42] A. Guo, H. Gu, J. Zhou, D. Mulhern, Y. Wang, K. A. Lee, V. Yang, M. Aguiar, J. Kornhauser, X. Jia, J. Ren, S. A. Beausoleil, J. C. Silva, V. Vemulapalli, M. T. Bedford, and M. J. Comb. Immunoaffinity enrichment and mass spectrometry analysis of protein methylation. *Mol. Cell Proteomics*, 13(1):372–387, Jan 2014. [PubMed Central:PMC3879628] [DOI:10.1074/mcp.O113.027870] [PubMed:24129315].

[43] H. Haase, T. Podzuweit, G. Lutsch, A. Hohaus, S. Kostka, C. Lindschau, M. Kott, R. Kraft, and I. Morano. Signaling from beta-adrenoceptor to L-type calcium channel: identification of a novel cardiac protein kinase A target possessing similarities to AHNAK. *FASEB J.*, 13(15):2161–2172, Dec 1999. [PubMed:10593863].

[44] R. Hamamoto, Y. Furukawa, M. Morita, Y. Iimura, F. P. Silva, M. Li, R. Yagyu, and Y. Nakamura. SMYD3 encodes a histone methyltransferase involved in the proliferation of cancer cells. *Nat. Cell Biol.*, 6(8):731–740, Aug 2004. [DOI:10.1038/ncb1151] [PubMed:15235609].

[45] R. Hamamoto, V. Saloura, and Y. Nakamura. Critical roles of non-histone protein lysine methylation in human tumorigenesis. *Nat. Rev. Cancer*, 15(2):110–124, Feb 2015. [DOI:10.1038/nrc3884] [PubMed:25614009].

[46] R. Hamamoto, F. P. Silva, M. Tsuge, T. Nishidate, T. Katagiri, Y. Nakamura, and Y. Furukawa. Enhanced SMYD3 expression is essential for the growth of breast cancer cells. *Cancer Sci.*, 97(2):113–118, Feb 2006. [DOI:10.1111/j.1349-7006.2006.00146.x] [PubMed:16441421].

[47] R. Hamamoto, G. Toyokawa, M. Nakakido, K. Ueda, and Y. Nakamura. SMYD2-dependent HSP90 methylation promotes cancer cell proliferation by regulating the chaperone complex formation. *Cancer Lett.*, 351(1):126–133, Aug 2014. [DOI:10.1016/j.canlet.2014.05.014] [PubMed:24880080].

[48] D. Hanahan and R. A. Weinberg. The hallmarks of cancer. *Cell*, 100(1):57–70, Jan 2000. [PubMed:10647931].

[49] D. Hanahan and R. A. Weinberg. Hallmarks of cancer: the next generation. *Cell*, 144(5):646–674, Mar 2011. [DOI:10.1016/j.cell.2011.02.013] [PubMed:21376230].

[50] T. Hashimoto, M. Amagai, D. A. Parry, T. W. Dixon, S. Tsukita, S. Tsukita, K. Miki, K. Sakai, Y. Inokuchi, and J. Kudoh. Desmoyokin, a 680 kDa keratinocyte plasma membrane-associated protein, is homologous to the protein encoded by human gene AHNAK. *J. Cell. Sci.*, 105 (Pt 2):275–286, Jun 1993. [PubMed:8408266].

[51] Y. He, I. Korboukh, J. Jin, and J. Huang. Targeting protein lysine methylation and demethylation in cancers. *Acta Biochim. Biophys. Sin. (Shanghai)*, 44(1):70–79, Jan 2012. [PubMed Central:PMC3244655] [DOI:10.1093/abbs/gmr109] [PubMed:22194015].

[52] D. Hu, X. Gao, M. A. Morgan, H. M. Herz, E. R. Smith, and A. Shilatifard. The MLL3/MLL4 branches of the COMPASS family function as major histone H3K4 monomethylases at enhancers. *Mol. Cell. Biol.*, 33(23):4745–4754, Dec 2013. [PubMed Central:PMC3838007] [DOI:10.1128/MCB.01181-13] [PubMed:24081332].

[53] L. Hu, Y. T. Zhu, C. Qi, and Y. J. Zhu. Identification of Smyd4 as a potential tumor suppressor gene involved in breast cancer development. *Cancer Res.*, 69(9):4067–4072, May 2009. [DOI:10.1158/0008-5472.CAN-08-4097] [PubMed:19383909].

[54] J. Huang and S. L. Berger. The emerging field of dynamic lysine methylation of non-histone proteins. *Curr. Opin. Genet. Dev.*, 18(2):152–158, Apr 2008. [DOI:10.1016/j.gde.2008.01.012] [PubMed:18339539].

[55] J. Huang, L. Perez-Burgos, B. J. Placek, R. Sengupta, M. Richter, J. A. Dorsey, S. Kubicek, S. Opravil, T. Jenuwein, and S. L. Berger. Repression of p53 activity by Smyd2-mediated methylation. *Nature*, 444(7119):629–632, Nov 2006. [DOI:10.1038/nature05287] [PubMed:17108971].

[56] NIH National Cancer Institute. *Cancer Statistics*, 2016. `https://www.cancer.gov/about-cancer/understanding/statistics` [Updated: March 14, 2016].

[57] E. Jacinto, R. Loewith, A. Schmidt, S. Lin, M. A. Ruegg, A. Hall, and M. N. Hall. Mammalian TOR complex 2 controls the actin cytoskeleton and is rapamycin insensitive. *Nat. Cell Biol.*, 6(11):1122–1128, Nov 2004. [DOI:10.1038/ncb1183] [PubMed:15467718].

[58] T. Jenuwein, G. Laible, R. Dorn, and G. Reuter. SET domain proteins modulate chromatin domains in eu- and heterochromatin. *Cell. Mol. Life Sci.*, 54(1):80–93, Jan 1998. [PubMed:9487389].

[59] Y. Jiang, N. Sirinupong, J. Brunzelle, and Z. Yang. Crystal structures of histone and p53 methyltransferase SmyD2 reveal a conformational flexibility of the autoinhibitory C-terminal domain. *PLoS ONE*, 6(6):e21640, 2011. [PubMed Central:PMC3125274] [DOI:10.1371/journal.pone.0021640] [PubMed:21738746].

[60] Y. Jiang, L. Trescott, J. Holcomb, X. Zhang, J. Brunzelle, N. Sirinupong, X. Shi, and Z. Yang. Structural insights into estrogen receptor Î± methylation by histone methyltransferase SMYD2, a cellular event implicated in estrogen signaling regulation. *J. Mol. Biol.*, 426(20):3413–3425, Oct 2014. [DOI:10.1016/j.jmb.2014.02.019] [PubMed:24594358].

[61] S. Just, B. Meder, I. M. Berger, C. Etard, N. Trano, E. Patzel, D. Hassel, S. Marquart, T. Dahme, B. Vogel, M. C. Fishman, H. A. Katus, U. Strahle, and W. Rottbauer. The myosin-interacting protein SMYD1 is essential for sarcomere organization. *J. Cell. Sci.*, 124(Pt 18):3127–3136, Sep 2011. [DOI:10.1242/jcs.084772] [PubMed:21852424].

[62] I. Kameshita, Z. Matsuda, T. Taniguchi, and Y. Shizuta. Poly (ADP-Ribose) synthetase. Separation and identification of three proteolytic fragments as the substrate-binding domain, the DNA-binding domain, and the automodification domain. *J. Biol. Chem.*, 259(8):4770–4776, Apr 1984. [PubMed:6325408].

[63] J. Kao, K. Salari, M. Bocanegra, Y. L. Choi, L. Girard, J. Gandhi, K. A. Kwei, T. Hernandez-Boussard, P. Wang, A. F. Gazdar, J. D. Minna, and J. R. Pollack. Molecular profiling of breast cancer cell lines defines relevant tumor models and provides a resource for cancer gene discovery. *PLoS ONE*, 4(7):e6146, Jul 2009. [PubMed Central:PMC2702084] [DOI:10.1371/journal.pone.0006146] [PubMed:19582160].

[64] F. Katzen. Gateway(®) recombinational cloning: a biological operating system. *Expert Opin Drug Discov*, 2(4):571–589, Apr 2007. [DOI:10.1517/17460441.2.4.571] [PubMed:23484762].

[65] S. Kawamura, E. Yoshigai, S. Kuhara, and K. Tashiro. smyd1 and smyd2 are expressed in muscle tissue in Xenopus laevis. *Cytotechnology*, 57(2):161–168, Jun 2008. [PubMed Central:PMC2553668] [DOI:10.1007/s10616-008-9128-1] [PubMed:19003161].

[66] M. C. Keogh, S. K. Kurdistani, S. A. Morris, S. H. Ahn, V. Podolny, S. R. Collins, M. Schuldiner, K. Chin, T. Punna, N. J. Thompson, C. Boone, A. Emili, J. S. Weissman, T. R. Hughes, B. D. Strahl, M. Grunstein, J. F. Greenblatt, S. Buratowski, and N. J. Krogan. Cotranscriptional set2 methylation of histone H3 lysine 36 recruits a repressive Rpd3 complex. *Cell*, 123(4):593–605, Nov 2005. [DOI:10.1016/j.cell.2005.10.025] [PubMed:16286008].

[67] P. D. Kingsley, K. E. McGrath, K. M. Maltby, A. D. Koniski, R. Ramchandran, and J. Palis. Subtractive hybridization reveals tissue-specific expression of ahnak during embryonic development. *Dev. Growth Differ.*, 43(2):133–143, Apr 2001. [PubMed:11284963].

[68] H. L. Ko and E. C. Ren. Functional Aspects of PARP1 in DNA Repair and Transcription. *Biomolecules*, 2(4):524–548, Nov 2012. [PubMed Central:PMC4030864] [DOI:10.3390/biom2040524] [PubMed:24970148].

[69] S. Komatsu, D. Ichikawa, S. Hirajima, H. Nagata, Y. Nishimura, T. Kawaguchi, M. Miyamae, W. Okajima, T. Ohashi, H. Konishi, A. Shiozaki, H. Fujiwara, K. Okamoto, H. Tsuda, I. Imoto, J. Inazawa, and E. Otsuji. Overexpression of SMYD2 contributes to malignant outcome in gastric cancer. *Br. J. Cancer*, 112(2):357–364, Jan 2015. [PubMed Central:PMC4453442] [DOI:10.1038/bjc.2014.543] [PubMed:25321194].

[70] S. Komatsu, I. Imoto, H. Tsuda, K. I. Kozaki, T. Muramatsu, Y. Shimada, S. Aiko, Y. Yoshizumi, D. Ichikawa, E. Otsuji, and J. Inazawa. Overexpression of SMYD2 relates to tumor cell proliferation and malignant outcome of esophageal squamous cell carcinoma. *Carcinogenesis*, 30(7):1139–1146, Jul 2009. [DOI:10.1093/carcin/bgp116] [PubMed:19423649].

[71] A. Komuro, Y. Masuda, K. Kobayashi, R. Babbitt, M. Gunel, R. A. Flavell, and V. T. Marchesi. The AHNAKs are a class of giant propeller-like proteins that associate with calcium channel proteins of cardiomyocytes and other cells. *Proc. Natl. Acad. Sci. U.S.A.*, 101(12):4053–4058, Mar 2004. [PubMed Central:PMC384694] [DOI:10.1073/pnas.0308619101] [PubMed:15007166].

[72] M. Kouno, G. Kondoh, K. Horie, N. Komazawa, N. Ishii, Y. Takahashi, J. Takeda, and T. Hashimoto. Ahnak/Desmoyokin is dispensable for proliferation, differentiation, and maintenance of integrity in mouse epidermis. *J. Invest. Dermatol.*, 123(4):700–707, Oct 2004. [DOI:10.1111/j.0022-202X.2004.23412.x] [PubMed:15373775].

[73] A. J. Kuo, P. Cheung, K. Chen, B. M. Zee, M. Kioi, J. Lauring, Y. Xi, B. H. Park, X. Shi, B. A. Garcia, W. Li, and O. Gozani. NSD2 links dimethylation of histone H3 at lysine 36 to oncogenic programming. *Mol. Cell*, 44(4):609–620, Nov 2011. [PubMed Central:PMC3222870] [DOI:10.1016/j.molcel.2011.08.042] [PubMed:22099308].

[74] S. Lanouette, J. A. Davey, F. Elisma, Z. Ning, D. Figeys, R. A. Chica, and J. F. Couture. Discovery of substrates for a SET domain lysine methyltransferase predicted by multistate computational protein design. *Structure*, 23(1):206–215, Jan 2015. [DOI:10.1016/j.str.2014.11.004] [PubMed:25533488].

[75] S. Lanouette, V. Mongeon, D. Figeys, and J. F. Couture. The functional diversity of protein lysine methylation. *Mol. Syst. Biol.*, 10:724, Apr 2014. [PubMed Central:PMC4023394] [PubMed:24714364].

[76] H. J. Lee and J. J. Zheng. PDZ domains and their binding partners: structure, specificity, and modification. *Cell Commun. Signal*, 8:8, May 2010. [PubMed Central:PMC2891790] [DOI:10.1186/1478-811X-8-8] [PubMed:20509869].

[77] I. H. Lee, H. J. Lim, S. Yoon, J. K. Seong, D. S. Bae, S. G. Rhee, and Y. S. Bae. Ahnak protein activates protein kinase C (PKC) through dissociation of the PKC-protein phosphatase 2A complex. *J. Biol. Chem.*, 283(10):6312–6320, Mar 2008. [DOI:10.1074/jbc.M706878200] [PubMed:18174170].

[78] I. H. Lee, J. O. You, K. S. Ha, D. S. Bae, P. G. Suh, S. G. Rhee, and Y. S. Bae. AHNAK-mediated activation of phospholipase C-gamma1 through protein kinase C. *J. Biol. Chem.*, 279(25):26645–26653, Jun 2004. [DOI:10.1074/jbc.M311525200] [PubMed:15033986].

[79] H. Li, J. Xu, Y. H. Bian, P. Rotllant, T. Shen, W. Chu, J. Zhang, M. Schneider, and S. J. Du. $Smyd1b_{t}v1, a key regulator of sarcomere assembly, is localized on the M-line of skeletal muscle fibers. PLoSONE, 6(12) : e28524, 2011. [PubMedCentral : PMC3235123][DOI : 10.1371/journal.pone.0028524][PubMed : 22174829]$.

[80] A. Lin, A. Minden, H. Martinetto, F. X. Claret, C. Lange-Carter, F. Mercurio, G. L. Johnson, and M. Karin. Identification of a dual specificity kinase that activates the Jun kinases and p38-Mpk2. *Science*, 268(5208):286–290, Apr 1995. [PubMed:7716521].

[81] L. Liu, X. T. Zhen, E. Denton, B. D. Marsden, and M. Schapira. ChromoHub: a data hub for navigators of chromatin-mediated signalling. *Bioinformatics*, 28(16):2205–2206, Aug 2012. [PubMed Central:PMC3413389] [DOI:10.1093/bioinformatics/bts340] [PubMed:22718786].

[82] Z. Liu, Y. Wang, T. Gao, Z. Pan, H. Cheng, Q. Yang, Z. Cheng, A. Guo, J. Ren, and Y. Xue. CPLM: a database of protein lysine modifications. *Nucleic Acids Res.*, 42(Database issue):D531–536, Jan 2014. [PubMed Central:PMC3964993] [DOI:10.1093/nar/gkt1093] [PubMed:24214993].

[83] K. J. Livak and T. D. Schmittgen. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods*, 25(4):402–408, Dec 2001. [DOI:10.1006/meth.2001.1262] [PubMed:11846609].

[84] R. Lu and G. G. Wang. Tudor: a versatile family of histone methylation 'readers'. *Trends Biochem. Sci.*, 38(11):546–555, Nov 2013. [PubMed Central:PMC3830939] [DOI:10.1016/j.tibs.2013.08.002] [PubMed:24035451].

[85] A. Marg, H. Haase, T. Neumann, M. Kouno, and I. Morano. AHNAK1 and AHNAK2 are costameric proteins: AHNAK1 affects transverse skeletal muscle fiber stiffness. *Biochem. Biophys. Res. Commun.*, 401(1):143–148, Oct 2010.

[86] P. K. Mazur, N. Reynoird, P. Khatri, P. W. Jansen, A. W. Wilkinson, S. Liu, O. Barbash, G. S. Van Aller, M. Huddleston, D. Dhanak, P. J. Tummino, R. G. Kruger, B. A. Garcia, A. J. Butte, M. Vermeulen, J. Sage, and O. Gozani. SMYD3 links lysine methylation of MAP3K2 to Ras-driven cancer. *Nature*, 510(7504):283–287, Jun 2014. [PubMed Central:PMC4122675] [DOI:10.1038/nature13320] [PubMed:24847881].

[87] T. Miller, N. J. Krogan, J. Dover, H. Erdjument-Bromage, P. Tempst, M. Johnston, J. F. Greenblatt, and A. Shilatifard. COMPASS: a complex of proteins associated with a trithorax-related SET domain protein. *Proc. Natl. Acad. Sci. U.S.A.*, 98(23):12902–12907, Nov 2001. [PubMed Central:PMC60797] [DOI:10.1073/pnas.231473398] [PubMed:11687631].

[88] L. Morera, M. Lubbert, and M. Jung. Targeting histone methyltransferases and demethylases in clinical trials for cancer therapy. *Clin Epigenetics*, 8:57, 2016. [PubMed Central:PMC4877953] [DOI:10.1186/s13148-016-0223-4] [PubMed:27222667].

[89] K. MURRAY. THE OCCURRENCE OF EPSILON-N-METHYL LYSINE IN HISTONES. *Biochemistry*, 3:10–15, Jan 1964. [PubMed:14114491].

[90] H. Nagandla, S. Lopez, W. Yu, T. L. Rasmussen, H. O. Tucker, R. J. Schwartz, and M. D. Stewart. Defective myogenesis in the absence of the muscle-specific lysine methyltransferase SMYD1. *Dev. Biol.*, 410(1):86–97, Feb 2016. [DOI:10.1016/j.ydbio.2015.12.005] [PubMed:26688546].

[91] M. Nakakido, Z. Deng, T. Suzuki, N. Dohmae, Y. Nakamura, and R. Hamamoto. Dysregulation of AKT Pathway by SMYD2-Mediated Lysine Methylation on PTEN. *Neoplasia*, 17(4):367–373, Apr 2015. [PubMed Central:PMC4415136] [DOI:10.1016/j.neo.2015.03.002] [PubMed:25925379].

[92] A. T. Nguyen and Y. Zhang. The diverse functions of Dot1 and H3K79 methylation. *Genes Dev.*, 25(13):1345–1358, Jul 2011. [PubMed Central:PMC3134078] [DOI:10.1101/gad.2057811] [PubMed:21724828].

[93] H. Nguyen, A. Allali-Hassani, S. Antonysamy, S. Chang, L. H. Chen, C. Curtis, S. Emtage, L. Fan, T. Gheyi, F. Li, S. Liu, J. R. Martin, D. Mendel, J. B. Olsen, L. Pelletier, T. Shatseva, S. Wu, F. F. Zhang, C. H. Arrowsmith, P. J. Brown, R. M. Campbell, B. A. Garcia, D. Barsyte-Lovejoy, M. Mader, and M. Vedadi. LLY-507, a Cell-active, Potent, and Selective Inhibitor of Protein-lysine Methyltransferase SMYD2. *J. Biol. Chem.*, 290(22):13641–13653, May 2015. [PubMed Central:PMC4447944] [DOI:10.1074/jbc.M114.626861] [PubMed:25825497].

[94] N. Ogata, K. Ueda, M. Kawaichi, and O. Hayaishi. Poly(ADP-ribose) synthetase, a main acceptor of poly(ADP-ribose) in isolated nuclei. *J. Biol. Chem.*, 256(9):4135–4137, May 1981. [PubMed:6260786].

[95] R. Ohtomo-Oda, S. Komatsu, T. Mori, S. Sekine, S. Hirajima, S. Yoshimoto, Y. Kanai, E. Otsuji, E. Ikeda, and H. Tsuda. SMYD2 overexpression is associated with tumor cell proliferation and a worse outcome in human papillomavirus-unrelated nonmultiple head and neck carcinomas. *Hum. Pathol.*, 49:145–155, Mar 2016. [DOI:10.1016/j.humpath.2015.08.025] [PubMed:26826421].

[96] J. B. Olsen, X. J. Cao, B. Han, L. H. Chen, A. Horvath, T. I. Richardson, R. M. Campbell, B. A. Garcia, and H. Nguyen. Quantitative Profiling of the Activity of Protein Lysine Methyltransferase SMYD2 Using SILAC-Based Proteomics. *Mol. Cell Proteomics*, 15(3):892–905, Mar 2016. [PubMed Central:PMC4813708] [DOI:10.1074/mcp.M115.053280] [PubMed:26750096].

[97] S. E. Ong, G. Mittler, and M. Mann. Identifying and quantifying in vivo methylation sites by heavy methyl SILAC. *Nat. Methods*, 1(2):119–126, Nov 2004. [DOI:10.1038/nmeth715] [PubMed:15782174].

[98] L. Piao, D. Kang, T. Suzuki, A. Masuda, N. Dohmae, Y. Nakamura, and R. Hamamoto. The histone methyltransferase SMYD2 methylates PARP1 and promotes poly(ADP-ribosyl)ation activity in cancer cells. *Neoplasia*, 16(3):257–264, Mar 2014. [PubMed Central:PMC4094795] [DOI:10.1016/j.neo.2014.03.002] [PubMed:24726141].

[99] I. A. Prior, P. D. Lewis, and C. Mattos. A comprehensive survey of Ras mutations in cancer. *Cancer Res.*, 72(10):2457–2467, May 2012. [PubMed Central:PMC3354961] [DOI:10.1158/0008-5472.CAN-11-2612] [PubMed:22589270].

[100] S. Qin and J. Min. Structure and function of the nucleosome-binding PWWP domain. *Trends Biochem. Sci.*, 39(11):536–547, Nov 2014. [DOI:10.1016/j.tibs.2014.09.001] [PubMed:25277115].

[101] K. Raj and G. J. Mufti. Azacytidine (Vidaza(R)) in the treatment of myelodysplastic syndromes. *Ther Clin Risk Manag*, 2(4):377–388, Dec 2006. [PubMed Central:PMC1936359] [PubMed:18360650].

[102] S. Rajan, H. Chu Pham Dang, H. Djambazian, H. Zuzan, Y. Fedyshyn, T. Ketela, J. Moffat, T. J. Hudson, and R. Sladek. Analysis of early C2C12 myogenesis identifies stably and differentially expressed transcriptional regulators whose knockdown inhibits myoblast differentiation. *Physiol. Genomics*, 44(2):183–197, Feb 2012. [DOI:10.1152/physiolgenomics.00093.2011] [PubMed:22147266].

[103] N. Reynoird, P. K. Mazur, T. Stellfeld, N. M. Flores, S. M. Lofgren, S. M. Carlson, E. Brambilla, P. Hainaut, E. B. Kaznowska, C. H. Arrowsmith, P. Khatri, C. Stresemann, O. Gozani, and J. Sage. Coordination of stress signals by the lysine methyltransferase SMYD2 promotes pancreatic cancer. *Genes Dev.*, 30(7):772–785, Apr 2016. [PubMed Central:PMC4826394] [DOI:10.1101/gad.275529.115] [PubMed:26988419].

[104] T. Rudolph, S. Beuch, and G. Reuter. Lysine-specific histone demethylase LSD1 and the dynamic control of chromatin. *Biol. Chem.*, 394(8):1019–1028, Aug 2013. [DOI:10.1515/hsz-2013-0119] [PubMed:23612539].

[105] L. A. Saddic, L. E. West, A. Aslanian, J. R. Yates, S. M. Rubin, O. Gozani, and J. Sage. Methylation of the retinoblastoma tumor suppressor by SMYD2. *J. Biol. Chem.*, 285(48):37733–37740, Nov 2010. [PubMed Central:PMC2988378] [DOI:10.1074/jbc.M110.137612] [PubMed:20870719].

[106] A. Sajjad, T. Novoyatleva, S. Vergarajauregui, C. Troidl, R. T. Schermuly, H. O. Tucker, and F. B. Engel. Lysine methyltransferase Smyd2 suppresses p53-dependent cardiomyocyte apoptosis. *Biochim. Biophys. Acta*, 1843(11):2556–2562, Nov 2014. [PubMed Central:PMC4157957] [DOI:10.1016/j.bbamcr.2014.06.019] [PubMed:25014164].

[107] L. H. Sakamoto, R. V. Andrade, M. S. Felipe, A. B. Motoyama, and F. Pittella Silva. SMYD2 is highly expressed in pediatric acute lymphoblastic leukemia and constitutes a bad prognostic factor. *Leuk. Res.*, 38(4):496–502, Apr 2014. [DOI:10.1016/j.leukres.2014.01.013] [PubMed:24631370].

[108] G. Sala, F. Dituri, C. Raimondi, S. Previdi, T. Maffucci, M. Mazzoletti, C. Rossi, M. Iezzi, R. Lattanzio, M. Piantelli, S. Iacobelli, M. Broggini, and M. Falasca. Phospholipase Cgamma1 is required for metastasis development and progression. *Cancer Res.*, 68(24):10187–10196, Dec 2008. [DOI:10.1158/0008-5472.CAN-08-1181] [PubMed:19074886].

[109] M. Schapira. Structural Chemistry of Human SET Domain Protein Methyltransferases. *Curr Chem Genomics*, 5(Suppl 1):85–94, 2011. [PubMed Central:PMC3178901] [DOI:10.2174/1875397301005010085] [PubMed:21966348].

[110] R. Schmieder, F. Puehler, R. Neuhaus, M. Kissel, A. A. Adjei, J. N. Miner, D. Mumberg, K. Ziegelbauer, and A. Scholz. Allosteric MEK1/2 inhibitor refametinib (BAY 86-9766) in combination with sorafenib exhibits antitumor activity in preclinical murine and rat models of hepatocellular carcinoma. *Neoplasia*, 15(10):1161–1171, Oct 2013. [PubMed Central:PMC3819632] [PubMed:24204195].

[111] F. Sekiya, Y. S. Bae, D. Y. Jhon, S. C. Hwang, and S. G. Rhee. AHNAK, a protein that binds and activates phospholipase C-gamma1 in the presence of arachidonic acid. *J. Biol. Chem.*, 274(20):13900–13907, May 1999. [PubMed:10318799].

[112] B. Sen, Z. Xie, N. Case, W. R. Thompson, G. Uzer, M. Styner, and J. Rubin. mTORC2 regulates mechanically induced cytoskeletal reorganization and lineage selection in marrow-derived mesenchymal stem cells. *J. Bone Miner. Res.*, 29(1):78–89, Jan 2014. [PubMed Central:PMC3870029] [DOI:10.1002/jbmr.2031] [PubMed:23821483].

[113] B. Sese, M. J. Barrero, M. C. Fabregat, V. Sander, and J. C. Izpisua Belmonte. SMYD2 is induced during cell differentiation and participates in early development. *Int. J. Dev. Biol.*, 57(5):357–364, 2013. [DOI:10.1387/ijdb.130051ji] [PubMed:23873367].

[114] J. Shankar, A. Messenberg, J. Chan, T. M. Underhill, L. J. Foster, and I. R. Nabi. Pseudopodial actin dynamics control epithelial-mesenchymal transition in metastatic cancer cells. *Cancer Res.*, 70(9):3780–3790, May 2010. [DOI:10.1158/0008-5472.CAN-09-4439] [PubMed:20388789].

[115] X. Shen, J. M. Collier, M. Hlaing, L. Zhang, E. H. Delshad, J. Bristow, and H. S. Bernstein. Genome-wide examination of myoblast cell cycle withdrawal during differentiation. *Dev. Dyn.*, 226(1):128–138, Jan 2003. [DOI:10.1002/dvdy.10200] [PubMed:12508234].

[116] Y. Shi, F. Lan, C. Matson, P. Mulligan, J. R. Whetstine, P. A. Cole, R. A. Casero, and Y. Shi. Histone demethylation mediated by the nuclear amine oxidase homolog LSD1. *Cell*, 119(7):941–953, Dec 2004. [DOI:10.1016/j.cell.2004.12.012] [PubMed:15620353].

[117] T. Shimazu, J. Barjau, Y. Sohtome, M. Sodeoka, and Y. Shinkai. Selenium-based S-adenosylmethionine analog reveals the mammalian seven-beta-strand methyltransferase METTL10 to be an EF1A1 lysine methyltransferase. *PLoS ONE*, 9(8):e105394, 2014. [PubMed Central:PMC4140779] [DOI:10.1371/journal.pone.0105394] [PubMed:25144183].

[118] E. Shtivelman and J. M. Bishop. The human gene AHNAK encodes a large phosphoprotein located primarily in the nucleus. *J. Cell Biol.*, 120(3):625–630, Feb 1993. [PubMed Central:PMC2119538] [PubMed:8381120].

[119] E. Shtivelman, F. E. Cohen, and J. M. Bishop. A human gene (AHNAK) encoding an unusually large protein with a 1.2-microns polyionic rod structure. *Proc. Natl. Acad. Sci. U.S.A.*, 89(12):5472–5476, Jun 1992. [PubMed Central:PMC49314] [PubMed:1608957].

[120] N. Sirinupong, J. Brunzelle, J. Ye, A. Pirzada, L. Nico, and Z. Yang. Crystal structure of cardiac-specific histone methyltransferase SmyD1 reveals unusual active site architecture. *J. Biol. Chem.*, 285(52):40635–40644, Dec 2010. [PubMed Central:PMC3003362] [DOI:10.1074/jbc.M110.168187] [PubMed:20943667].

[121] C. J. Sneeringer, M. P. Scott, K. W. Kuntz, S. K. Knutson, R. M. Pollock, V. M. Richon, and R. A. Copeland. Coordinated activities of wild-type plus mutant EZH2 drive tumor-associated hypertrimethylation of lysine 27 on histone H3 (H3K27) in human B-cell lymphomas. *Proc. Natl. Acad. Sci. U.S.A.*, 107(49):20980–20985, Dec 2010. [PubMed Central:PMC3000297] [DOI:10.1073/pnas.1012525107] [PubMed:21078963].

[122] R. Spadaccini, H. Perrin, M. J. Bottomley, S. Ansieau, and M. Sattler. Structure and functional analysis of the MYND domain. *J. Mol. Biol.*, 358(2):498–508, Apr 2006. [DOI:10.1016/j.jmb.2006.01.087] [PubMed:16527309].

[123] J. D. Stender, G. Pascual, W. Liu, M. U. Kaikkonen, K. Do, N. J. Spann, M. Boutros, N. Perrimon, M. G. Rosenfeld, and C. K. Glass. Control of proinflammatory gene programs by regulated trimethylation and demethylation of histone H4K20. *Mol. Cell*, 48(1):28–38, Oct 2012. [PubMed Central:PMC3472359] [DOI:10.1016/j.molcel.2012.07.020] [PubMed:22921934].

[124] B. W. Stewart and C. P. Wild. *World Cancer Report 2014*, volume 3. WHO Press, 2014.

[125] H. Sudo, A. B. Tsuji, A. Sugyo, M. Abe, O. Hino, and T. Saga. AHNAK is highly expressed and plays a key role in cell migration and invasion in mesothelioma. *Int. J. Oncol.*, 44(2):530–538, Feb 2014. [DOI:10.3892/ijo.2013.2183] [PubMed:24253341].

[126] T. Suganuma, S. G. Pattenden, and J. L. Workman. Diverse functions of WD40 repeat proteins in histone recognition. *Genes Dev.*, 22(10):1265–1268, May 2008. [PubMed Central:PMC2732410] [DOI:10.1101/gad.1676208] [PubMed:18483215].

[127] X. Tan, J. Rotllant, H. Li, P. De Deyne, P. DeDeyne, and S. J. Du. SmyD1, a histone methyltransferase, is required for myofibril organization and muscle contraction in zebrafish embryos. *Proc. Natl. Acad. Sci. U.S.A.*, 103(8):2713–2718, Feb 2006. [PubMed Central:PMC1531647] [DOI:10.1073/pnas.0509503103] [PubMed:16477022].

[128] E. C. Thompson and A. A. Travers. A Drosophila Smyd4 homologue is a muscle-specific transcriptional modulator involved in development. *PLoS ONE*, 3(8):e3008, Aug 2008. [PubMed Central:PMC2500188] [DOI:10.1371/journal.pone.0003008] [PubMed:18714374].

[129] B. Tschiersch, A. Hofmann, V. Krauss, R. Dorn, G. Korge, and G. Reuter. The protein encoded by the Drosophila position-effect variegation suppressor gene Su(var)3-9 combines domains of antagonistic regulators of homeotic gene complexes. *EMBO J.*, 13(16):3822–3831, Aug 1994. [PubMed Central:PMC395295] [PubMed:7915232].

[130] Y. Tsukada, J. Fang, H. Erdjument-Bromage, M. E. Warren, C. H. Borchers, P. Tempst, and Y. Zhang. Histone demethylation by a family of JmjC domain-containing proteins. *Nature*, 439(7078):811–816, Feb 2006. [DOI:10.1038/nature04433] [PubMed:16362057].

[131] T. Voelkel, C. Andresen, A. Unger, S. Just, W. Rottbauer, and W. A. Linke. Lysine methyltransferase Smyd2 regulates Hsp90-mediated protection of the sarcomeric titin springs and cardiac function. *Biochim. Biophys. Acta*, 1833(4):812–822, Apr 2013. [DOI:10.1016/j.bbamcr.2012.09.012] [PubMed:23047121].

[132] E. J. Wagner and P. B. Carpenter. Understanding the language of Lys36 methylation at histone H3. *Nat. Rev. Mol. Cell Biol.*, 13(2):115–126, Jan 2012. [PubMed Central:PMC3969746] [DOI:10.1038/nrm3274] [PubMed:22266761].

[133] L. Wang, L. Li, H. Zhang, X. Luo, J. Dai, S. Zhou, J. Gu, J. Zhu, P. Atadja, C. Lu, E. Li, and K. Zhao. Structure of human SMYD2 protein reveals the basis of p53 tumor suppressor methylation. *J. Biol. Chem.*, 286(44):38725–38737, Nov 2011. [PubMed Central:PMC3207477] [DOI:10.1074/jbc.M111.262410] [PubMed:21880715].

[134] R. A. Weinberg. *The Biology of Cancer*. Taylor & Francis, 2nd edition, 2013.

[135] R. P. Wernyj, C. M. Ewing, and W. B. Isaacs. Multiple antibodies to titin immunoreact with AHNAK and localize to the mitotic spindle machinery. *Cell Motil. Cytoskeleton*, 50(2):101–113, Oct 2001. [DOI:10.1002/cm.1044] [PubMed:11746675].

[136] H. Wu, J. Min, V. V. Lunin, T. Antoshenko, L. Dombrovski, H. Zeng, A. Allali-Hassani, V. Campagna-Slater, M. Vedadi, C. H. Arrowsmith, A. N. Plotnikov, and M. Schapira. Structural biology of human H3K9 methyltransferases. *PLoS ONE*, 5(1):e8570, Jan 2010. [PubMed Central:PMC2797608] [DOI:10.1371/journal.pone.0008570] [PubMed:20084102].

[137] J. Wu, T. Cheung, C. Grande, A. D. Ferguson, X. Zhu, K. Theriault, E. Code, C. Birr, N. Keen, and H. Chen. Biochemical characterization of human SET and MYND domain-containing protein 2 methyltransferase. *Biochemistry*, 50(29):6488–6497, Jul 2011. [DOI:10.1021/bi200725p] [PubMed:21678921].

[138] D. Yaffe and O. Saxel. Serial passaging and differentiation of myogenic cells isolated from dystrophic mouse muscle. *Nature*, 270(5639):725–727, 1977. [PubMed:563524].

[139] J. S. You and P. A. Jones. Cancer genetics and epigenetics: two sides of the same coin? *Cancer Cell*, 22(1):9–20, Jul 2012. [PubMed Central:PMC3396881] [DOI:10.1016/j.ccr.2012.06.008] [PubMed:22789535].

[140] J. Yuan and J. Chen. FIGNL1-containing protein complex is required for efficient homologous recombination repair. *Proc. Natl. Acad. Sci. U.S.A.*, 110(26):10640–10645, Jun 2013. [PubMed Central:PMC3696823] [DOI:10.1073/pnas.1220662110] [PubMed:23754376].

[141] X. Zhang, K. Tanaka, J. Yan, J. Li, D. Peng, Y. Jiang, Z. Yang, M. C. Barton, H. Wen, and X. Shi. Regulation of estrogen receptor Î± by histone methyltransferase SMYD2-mediated protein methylation. *Proc. Natl. Acad. Sci. U.S.A.*, 110(43):17284–17289, Oct 2013. [PubMed Central:PMC3808627] [DOI:10.1073/pnas.1307959110] [PubMed:24101509].

Affirmation/Eidesstattliche Versicherung

Hereby I declare that I submit the current PhD thesis for the first time and that all of the following material was written by myself and is properly referenced.

Hiermit versichere ich, dass ich die vorliegende Doktorarbeit erstmalig einreiche, selbständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel verwendet habe.

Ort, Datum:

Unterschrift:

Erik Eggert

7 Supplement

7.1 Translated protein sequences of expression constructs

1CRU-N-term. GST-fusion (red) protein incl. eTEV cleavage site (yellow)
(388 aa, theor. MW: 44.2 kDa, pI 6,04)
Represents aa 3987-4129 of human full-length AHNAK (Uniprot ID Q09666)

```
  1 MSPILGYWKI KGLVQPTRLL LEYLEEKYEE HLYERDEGDK WRNKKFELGL EFPNLPYYID
 61 GDVKLTQSMA IIRYIADKHN MLGGCPKERA EISMLEGAVL DIRYGVSRIA YSKDFETLKV
121 DFLSKLPEML KMFEDRLCHK TYLNGDHVTH PDFMLYDALD VVLYMDPMCL DAFPKLVCFK
181 KRIEAIPQID KYLKSSKYIA WPLQGWQATF GGGDHPPKSD PITSLYKKAG Fdydiptten
241 lyfqgPKFKM PEMNIKAPKI SMPDFDLHLK GPKVKGDVDV SLPKMEGDLK APEVDIKGPK
301 VDIDAPDVDV HGPDWHLKMP KVKMPKFSMP GFKGEGPEVD VNLPKADIDV SGPKVDIDTP
361 DIDIHGPEGK LKGPKFKMPD LHLKAPKI
```

3sCRU-N-term. GST-fusion (red) protein incl. eTEV cleavage site (yellow)
(453 aa, theor. MW: 51.0 kDa, pI 6.01)
Represents aa 4384-4591 of human full-length AHNAK (Uniprot ID Q09666)

```
  1 MSPILGYWKI KGLVQPTRLL LEYLEEKYEE HLYERDEGDK WRNKKFELGL EFPNLPYYID
 61 GDVKLTQSMA IIRYIADKHN MLGGCPKERA EISMLEGAVL DIRYGVSRIA YSKDFETLKV
121 DFLSKLPEML KMFEDRLCHK TYLNGDHVTH PDFMLYDALD VVLYMDPMCL DAFPKLVCFK
181 KRIEAIPQID KYLKSSKYIA WPLQGWQATF GGGDHPPKSD PITSLYKKAG Fdydiptten
241 lyfqgKISMP DIDFNLKGPK VKGDVDVSLP KVEGDLKGPE IDIKGPSLDI DTPDVNIEGP
301 EGKLKGPKFK MPEMNIKAPK ISMPDFDLHL KGPKVKGDVD VSLPKVESDL KGPEVDIEGP
361 EGKLKGPKFK MPDVHFKSPQ ISMSDIDLNL KGPKIKGDMD ISVPKLEGDL KGPKVDVKGP
421 KVGIDTPDID IHGPEGKLKG PKFKMPDLHL KAP
```

3lCRU-N-term. GST-fusion (red) protein incl. eTEV cleavage site (yellow)
(629 aa, theor. MW: 70.3 kDa, pI 5.54)
Represents aa 1066-1449 of human full-length AHNAK (Uniprot ID Q09666)

```
  1 MSPILGYWKI KGLVQPTRLL LEYLEEKYEE HLYERDEGDK WRNKKFELGL EFPNLPYYID
 61 GDVKLTQSMA IIRYIADKHN MLGGCPKERA EISMLEGAVL DIRYGVSRIA YSKDFETLKV
121 DFLSKLPEML KMFEDRLCHK TYLNGDHVTH PDFMLYDALD VVLYMDPMCL DAFPKLVCFK
181 KRIEAIPQID KYLKSSKYIA WPLQGWQATF GGGDHPPKSD PITSLYKKAG Fdydiptten
241 lyfqgKMSLP DVDLDLKGPK MKGNVDISAP KIEGEMQVPD VDIRGPKVDI KAPDVEGQGL
301 DWSLKIPKMK MPKFSMPSLK GEGPEVDVNL PKADVVVSGP KVDIEAPDVS LEGPEGKLKG
361 PKFKMPEMHF KTPKISMPDV DLHLKGPKVK GDVDVSVPKV EGEMKVPDVE IKGPKMDIDA
421 PDVEVQGPDW HLKMPKMKMP KFSMPGFKGE GREVDVNLPK ADIDVSGPKV DVEVPDVSLE
481 GPEGKLKGPK FKMPEMHFKA PKISMPDVDL NLKGPKLKGD VDVSLPEVEG EMKVPDVDIK
541 GPKVDISAPD VDVHGPDWHL KMPKVKMPKF SMPGFKGEGP EVDVKLPKAD VDVSGPKMDA
601 EVPDVNIEGP DAKLKGPKFK MPEMSIKPQ
```

Human SMYD2 with N-terminal 2x-c-myc tag (yellow)
(454 aa , theor. MW: 52.0 kDa, pI 5.80)
Histidine marked in red corresponds to endogenous SMYD2 (NP_064582) position H207 and was replaced in the enzymatic inactive mutant (H207A) by alanine.

```
  1 MEQKLISEED LEQKLISEED LRAEGLGGLE RFCSPGKGRG LRALQPFQVG DLLFSCPAYA
 61 YVLTVNERGN HCEYCFTRKE GLSKCGRCKQ AFYCNVECQK EDWPMHKLEC SPMVVFGENW
121 NPSETVRLTA RILAKQKIHP ERTPSEKLLA VKEFESHLDK LDNEKKDLIQ SDIAALHHFY
181 SKHLGFPDND SLVVLFAQVN CNGFTIEDEE LSHLGSAIFP DVALMNHSCC PNVIVTYKGT
241 LAEVRAVQEI KPGEEVFTSY IDLLYPTEDR NDRLRDSYFF TCECQECTTK DKDKAKVEIR
301 KLSDPPKAEA IRDMVRYARN VIEEFRRAKH YKSPSELLEI CELSQEKMSS VFEDSNVYML
361 HMMYQAMGVC LYMQDWEGAL QYGQKIIKPY SKHYPLYSLN VASMWLKLGR LYMGLEHKAA
421 GEKALKKAIA IMEVAHGKDH PYISEIKQEI ESH
```

Human SMYD3 isoform 1 with N-terminal 2x-FLAG tag (yellow)
(444 aa , theor. MW: 51.0 kDa, pI 5.96)
Asparagine marked in red corresponds to endogenous SMYD3 (NP_001161212) position N205 and was replaced in the enzymatic inactive mutant (N205A) by alanine.

```
  1 MDYKDDDDKD YKDDDDKEPL KVEKFATAKR GNGLRAVTPL RPGELLFRSD PLAYTVCKGS
 61 RGVVCDRCLL GKEKLMRCSQ CRVAKYCSAK CQKKAWPDHK RECKCLKSCK PRYPPDSVRL
121 LGRVVFKLMD GAPSESEKLY SFYDLESNIN KLTEDKKEGL RQLVMTFQHF MREEIQDASQ
181 LPPAFDLFEA FAKVICNSFT ICNAEMQEVG VGLYPSISLL NHSCDPNCSI VFNGPHLLLR
241 AVRDIEVGEE LTICYLDMLM TSEERRKQLR DQYCFECDCF RCQTQDKDAD MLTGDEQVWK
301 EVQESLKKIE ELKAHWKWEQ VLAMCQAIIS SNSERLPDIN IYQLKVLDCA MDACINLGLL
361 EEALFYGTRT MEPYRIFFPG SHPVRGVQVM KVGKLQLHQG MFPQAMKNLR LAFDIMRVTH
421 GREHSLIEDL ILLLEECDAN IRAS
```

eGFP
(240 aa , theor. MW: 26.9 kDa, pI 5.58)

```
  1 MVSKGEELFT GVVPILVELD GDVNGHKFSV SGEGEGDATY GKLTLKFICT TGKLPVPWPT
 61 LVTTLTYGVQ CFSRYPDHMK QHDFFKSAMP EGYVQERTIF FKDDGNYKTR AEVKFEGDTL
121 VNRIELKGID FKEDGNILGH KLEYNYNSHN VYIMADKQKN GIKVNFKIRH NIEDGSVQLA
181 DHYQQNTPIG DGPVLLPDNH YLSTQSALSK DPNEKRDHMV LLEFVTAAGI TLGMDELYK
```

7.2 List of mono-methylated peptides detected in proteom study

Tab. 7.10: Detected lysine mono-methylation sites

Antibody	Gene	Mass, kDa	Uniprot ID	Kme1 site
pan-Kme1	DST	861	Q03001	K590 [PLLK(me1)SSL]
pan-Kme1	DST	861	Q03001	K6757 [LGAK(me1)HSV]
pan-Kme1	MACF1	838	Q9UPN3	K6604 [LGGK(me1)QPV]
both	SYNE2	796	Q8WXH0-9	K265 [EAYK(me1)NF_]
both	AHNAK	629	Q09666	K1012 [PSLK(me1)GEG]
both	AHNAK	629	Q09666	K1024 [NLSK(me1)ANV]
pan-Kme1	AHNAK	629	Q09666	K1077 [LDLK(me1)GPK]
pan-Kme1	AHNAK	629	Q09666	K1080 [KGPK(me1)MKG]
pan-Kme1	AHNAK	629	Q09666	K1082 [PKMK(me1)GNV]
SY46	AHNAK	629	Q09666	K1184 [PKFK(me1)MPE]
both	AHNAK	629	Q09666	K1191 [MHFK(me1)TPK]
both	AHNAK	629	Q09666	K1205 [LHLK(me1)GPK]
pan-Kme1	AHNAK	629	Q09666	K1208 [KGPK(me1)VKG]
pan-Kme1	AHNAK	629	Q09666	K1268 [PGFK(me1)GEG]
both	AHNAK	629	Q09666	K1319 [MHFK(me1)APK]
both	AHNAK	629	Q09666	K1333 [LNLK(me1)GPK]
pan-Kme1	AHNAK	629	Q09666	K1336 [KGPK(me1)LKG]
pan-Kme1	AHNAK	629	Q09666	K1396 [PGFK(me1)GEG]
both	AHNAK	629	Q09666	K1461 [LHLK(me1)GPK]
both	AHNAK	629	Q09666	K1464 [KGPK(me1)MKG]
pan-Kme1	AHNAK	629	Q09666	K1524 [PGFK(me1)GEG]
pan-Kme1	AHNAK	629	Q09666	K1589 [LNLK(me1)APK]
pan-Kme1	AHNAK	629	Q09666	K1594 [PKLK(me1)TDV]
pan-Kme1	AHNAK	629	Q09666	K1791 [LNLK(me1)GPK]
pan-Kme1	AHNAK	629	Q09666	K1911 [PKFK(me1)MPD]
pan-Kme1	AHNAK	629	Q09666	K1918 [MHFK(me1)APK]
both	AHNAK	629	Q09666	K2194 [LNLK(me1)GPK]
both	AHNAK	629	Q09666	K2197 [KGPK(me1)VKG]
pan-Kme1	AHNAK	629	Q09666	K2257 [PGFK(me1)GEG]
both	AHNAK	629	Q09666	K2322 [FNLK(me1)GPK]
both	AHNAK	629	Q09666	K2396 [LHLK(me1)SPK]
pan-Kme1	AHNAK	629	Q09666	K243 [GHSK(me1)LQV]
pan-Kme1	AHNAK	629	Q09666	K2449 [MHFK(me1)APN]
pan-Kme1	AHNAK	629	Q09666	K2463 [LNLK(me1)GPK]
pan-Kme1	AHNAK	629	Q09666	K2466 [KGPK(me1)IKG]
both	AHNAK	629	Q09666	K2589 [LHLK(me1)GPK]
pan-Kme1	AHNAK	629	Q09666	K2652 [PGFK(me1)GEG]
pan-Kme1	AHNAK	629	Q09666	K2973 [LHLK(me1)GPK]
pan-Kme1	AHNAK	629	Q09666	K3164 [PGFK(me1)GEG]
both	AHNAK	629	Q09666	K3229 [LNLK(me1)GPK]
pan-Kme1	AHNAK	629	Q09666	K3249 [GDLK(me1)GPA]
pan-Kme1	AHNAK	629	Q09666	K3306 [KGSK(me1)LKG]
pan-Kme1	AHNAK	629	Q09666	K3404 [SKFK(me1)MPF]
pan-Kme1	AHNAK	629	Q09666	K3425 [LNLK(me1)SPK]

Tab. 7.10: Detected lysine mono-methylation sites

Antibody	Gene	Mass, kDa	Uniprot ID	Kme1 site
pan-Kme1	AHNAK	629	Q09666	K3527 [GGLK(me1)GPK]
both	AHNAK	629	Q09666	K3681 [LNLK(me1)GPK]
pan-Kme1	AHNAK	629	Q09666	K3684 [KGPK(me1)MKG]
pan-Kme1	AHNAK	629	Q09666	K3686 [PKMK(me1)GDV]
pan-Kme1	AHNAK	629	Q09666	K371 [GKLK(me1)GPQ]
both	AHNAK	629	Q09666	K3741 [MHLK(me1)APK]
both	AHNAK	629	Q09666	K4125 [LHLK(me1)APK]
pan-Kme1	AHNAK	629	Q09666	K4139 [LNLK(me1)GPK]
pan-Kme1	AHNAK	629	Q09666	K4142 [KGPK(me1)MKG]
pan-Kme1	AHNAK	629	Q09666	K4330 [PGFK(me1)GEG]
both	AHNAK	629	Q09666	K4395 [FNLK(me1)GPK]
both	AHNAK	629	Q09666	K4515 [VHFK(me1)SPQ]
both	AHNAK	629	Q09666	K4529 [LNLK(me1)GPK]
pan-Kme1	AHNAK	629	Q09666	K4532 [KGPK(me1)IKG]
pan-Kme1	AHNAK	629	Q09666	K4794 [PGFK(me1)GEG]
both	AHNAK	629	Q09666	K4859 [LDLK(me1)GPK]
pan-Kme1	AHNAK	629	Q09666	K4862 [KGPK(me1)VKG]
both	AHNAK	629	Q09666	K4879 [GTLK(me1)GPE]
both	AHNAK	629	Q09666	K4886 [VDLK(me1)GPR]
pan-Kme1	AHNAK	629	Q09666	K4905 [PSLK(me1)MPS]
pan-Kme1	AHNAK	629	Q09666	K4926 [LHLK(me1)APK]
both	AHNAK	629	Q09666	K4945 [VDLK(me1)GPK]
pan-Kme1	AHNAK	629	Q09666	K4985 [FGAK(me1)SPK]
pan-Kme1	AHNAK	629	Q09666	K5204 [VNLK(me1)GPK]
both	AHNAK	629	Q09666	K5281 [VSLK(me1)GPG]
both	AHNAK	629	Q09666	K5307 [LNLK(me1)GPS]
both	AHNAK	629	Q09666	K5312 [PSLK(me1)GDL]
SY46	AHNAK	629	Q09666	K5323 [PSMK(me1)VHA]
pan-Kme1	AHNAK	629	Q09666	K537 [VALK(me1)GSR]
both	AHNAK	629	Q09666	K5434 [GELK(me1)GPG]
both	AHNAK	629	Q09666	K5443 [VNLK(me1)GPR]
pan-Kme1	AHNAK	629	Q09666	K5528 [GGLK(me1)GSE]
both	AHNAK	629	Q09666	K5725 [PKGK(me1)GGV]
pan-Kme1	AHNAK	629	Q09666	K584 [PKVK(me1)GGV]
both	AHNAK	629	Q09666	K828 [LHLK(me1)GPN]
pan-Kme1	AHNAK	629	Q09666	K855 [VELK(me1)SAK]
pan-Kme1	AHNAK	629	Q09666	K891 [PGFK(me1)AEG]
pan-Kme1	AHNAK	629	Q09666	K942 [MNIK(me1)APK]
pan-Kme1	AHNAK	629	Q09666	K956 [LHMK(me1)GPK]
both	AHNAK2	617	Q8IVF2	K1097 [VALK(me1)GPQ]
both	AHNAK2	617	Q8IVF2	K1104 [VDVK(me1)GPK]
both	AHNAK2	617	Q8IVF2	K1427 [VDLK(me1)GPE]
both	AHNAK2	617	Q8IVF2	K1757 [VALK(me1)GPQ]
both	AHNAK2	617	Q8IVF2	K1764 [MDVK(me1)GPK]
pan-Kme1	AHNAK2	617	Q8IVF2	K2094 [VDIK(me1)GPK]
pan-Kme1	AHNAK2	617	Q8IVF2	K2582 [MDLK(me1)GPQ]

Tab. 7.10: Detected lysine mono-methylation sites

Antibody	Gene	Mass, kDa	Uniprot ID	Kme1 site
pan-Kme1	AHNAK2	617	Q8IVF2	K2589 [LDVK(me1)GPK]
pan-Kme1	AHNAK2	617	Q8IVF2	K3249 [IDIK(me1)GPK]
both	AHNAK2	617	Q8IVF2	K3407 [VDLK(me1)SPQ]
pan-Kme1	AHNAK2	617	Q8IVF2	K3414 [VDIK(me1)GPK]
pan-Kme1	AHNAK2	617	Q8IVF2	K3417 [KGPK(me1)LDL]
pan-Kme1	AHNAK2	617	Q8IVF2	K4397 [VDLK(me1)GPQ]
both	AHNAK2	617	Q8IVF2	K4562 [VDLK(me1)GPQ]
both	AHNAK2	617	Q8IVF2	K4569 [VDVK(me1)GPK]
pan-Kme1	AHNAK2	617	Q8IVF2	K839 [PKFK(me1)MPS]
both	AHNAK2	617	Q8IVF2	K932 [VDLK(me1)GPQ]
both	AHNAK2	617	Q8IVF2	K939 [IDVK(me1)GPK]
pan-Kme1	AHNAK2	617	Q8IVF2	K942 [KGPK(me1)LDL]
pan-Kme1	EPPK1	556	P58107	K680 [VFAK(me1)LLS]
pan-Kme1	DYNC1H1	532	Q14204	K3480 [ALLK(me1)SLS]
pan-Kme1	DYNC1H1	532	Q14204	K4154 [PGVK(me1)ANM]
pan-Kme1	PLEC	532	Q15149	K915 [NEYK(me1)GHL]
pan-Kme1	DYNC2H1	493	Q8NCM8	K2455 [KNLK(me1)NHS]
pan-Kme1	PRKDC	469	P78527	K2694 [APLK(me1)SVG]
pan-Kme1	VPS13B	449	Q7Z7G8	K2052 [KPLK(me1)ANL]
pan-Kme1	KMT2A	432	Q03164	K3804 [VQLK(me1)SAR]
SY46	USP34	404	Q70CQ2	K1201 [SHLK(me1)ALS]
pan-Kme1	SPEN	402	Q96T58	K2680 [TTLK(me1)SLV]
pan-Kme1	VPS13A	360	Q96RL7	K1454 [FSLK(me1)NCI]
pan-Kme1	HTT	348	P42858	K9 [KLMK(me1)AFE]
pan-Kme1	EP400	343	Q96L91	K3018 [PALK(me1)TQF]
pan-Kme1	BPTF	338	Q12830	K791 [SQLK(me1)SQQ]
pan-Kme1	BPTF	338	Q12830	K833 [VIMK(me1)GNI]
pan-Kme1	CRYBG3	331	Q68DQ2	K1193 [LDLK(me1)SSL]
pan-Kme1	CRYBG3	331	Q68DQ2	K293 [SSMK(me1)GNL]
pan-Kme1	BOD1L1	330	Q8NFC6	K123 [QVLK(me1)SGM]
pan-Kme1	LRBA	319	P50851-2	K212 [FPGK(me1)SAA]
both	PRRC2C	317	Q9Y520-7	K27 [NTYK(me1)GKS]
both	PRRC2C	317	Q9Y520-7	K29 [YKGK(me1)SLE]
pan-Kme1	PRRC2C	317	Q9Y520-7	K35 [ETQK(me1)TTV]
pan-Kme1	PRRC2C	317	Q9Y520-7	K437 [FPSK(me1)QQV]
pan-Kme1	ITPR1	314	Q14643	K1574 [LFLK(me1)SHS]
pan-Kme1	USP24	294	Q9UPU5	K620 [DGFK(me1)SSQ]
both	GCN1L1	293	Q92616	K1663 [PGLK(me1)ASL]
pan-Kme1	GCN1L1	293	Q92616	K802 [YSFK(me1)EQI]
pan-Kme1	FLNC	291	Q14315	K1073 [PGLK(me1)GGL]
pan-Kme1	MTOR	289	P42345	K2496 [LNKK(me1)AIQ]
both	MAST4	284	O15021	K1868 [SSLK(me1)MNK]
both	MAST4	284	O15021	K1871 [KMNK(me1)SYL]
SY46	SPTA1	280	P02549	K1486 [KALK(me1)AQL]
pan-Kme1	SPG11	279	Q96JI7	K2200 [GTLK(me1)TAL]
pan-Kme1	SPG11	279	Q96JI7	K2374 [RLLK(me1)SSI]

Tab. 7.10: Detected lysine mono-methylation sites

Antibody	Gene	Mass, kDa	Uniprot ID	Kme1 site
pan-Kme1	FLNB	278	O75369	K1339 [PGLK(me1)EAF]
both	FLNB	278	O75369-8	K1477 [SPLK(me1)ALS]
SY46	DOPEY1	277	Q5JWR5	K1273 [FSFK(me1)EKL]
SY46	DOPEY1	277	Q5JWR5	K1275 [FKEK(me1)LSE]
SY46	DOPEY1	277	Q5JWR5	K414 [ENKK(me1)TAE]
pan-Kme1	BAHCC1	277	Q9P281	K659 [GGLK(me1)ASC]
pan-Kme1	BAHCC1	277	Q9P281	K713 [ARQK(me1)DTV]
SY46	SPTBN1	275	Q01082	K1156 [ELHK(me1)MWE]
pan-Kme1	ANKRD17	274	O75179	K879 [LQLK(me1)TQQ]
pan-Kme1	PRPF8	274	Q6P2Q9	K1443 [TDFK(me1)QYQ]
pan-Kme1	TLN1	270	Q9Y490	K1332 [PNLK(me1)SQL]
pan-Kme1	TLN1	270	Q9Y490	K2266 [PELK(me1)QQL]
pan-Kme1	NAV2	268	Q8IVL1	K59 [IPLK(me1)SQV]
pan-Kme1	CASC5	265	Q8NG31	K846 [GVLK(me1)SNC]
pan-Kme1	AHCTF1	252	Q8WYP5	K969 [PALK(me1)LNQ]
pan-Kme1	ASCC3	251	Q8N3C0	K1382 [APLK(me1)ALV]
pan-Kme1	ASCC3	251	Q8N3C0	K938 [ISHK(me1)AYQ]
pan-Kme1	RTTN	249	Q86VV8	K1882 [HALK(me1)ANL]
pan-Kme1	SVIL	248	O95425	K927 [SILK(me1)SQA]
pan-Kme1	CABIN1	246	Q9Y6J0	K2173 [QKLK(me1)SAI]
pan-Kme1	CAD	243	P27708	K1301 [AYLK(me1)AML]
pan-Kme1	CAD	243	P27708	K186 [CGLK(me1)YNQ]
pan-Kme1	CAD	243	P27708	K1963 [LKGK(me1)VMA]
pan-Kme1	MYOF	235	Q9NZM1	K1416 [PQLK(me1)ASL]
pan-Kme1	FANCM	232	Q8IYD8	K1343 [PLSK(me1)SNT]
pan-Kme1	RALGAPA1	230	Q6GYQ0	K1133 [WLFK(me1)ATM]
pan-Kme1	MYH9	227	P35579	K1918 [SSLK(me1)NKL]
pan-Kme1	MYH9	227	P35579	K1920 [LKNK(me1)LRR]
pan-Kme1	DOCK4	225	Q8N1I0	K1818 [SPLK(me1)GSV]
pan-Kme1	MICAL3	224	Q7RTP6	K1364 [VSLK(me1)SYS]
pan-Kme1	HEATR5A	222	Q86XA9	K2036 [IQLK(me1)TSF]
pan-Kme1	TANC2	220	Q9HCD6	K7 [NSLK(me1)MLL]
pan-Kme1	ARHGAP21	217	Q5T5U3	K346 [ILLK(me1)SGN]
pan-Kme1	ARHGAP21	217	Q5T5U3	K590 [PDLK(me1)TLQ]
pan-Kme1	ARHGAP21	217	Q5T5U3	K897 [LSFK(me1)HVS]
pan-Kme1	POLR2A	217	P24928	K767 [NNFK(me1)SMV]
SY46	ALG9	215	Q9H6U8	K11 [QRLK(me1)GSG]
pan-Kme1	DOCK5	215	Q9H7D0	K1765 [QRPK(me1)SLQ]
pan-Kme1	KIAA1217	214	Q5T5P2	K1872 [HTGK(me1)GHH]
pan-Kme1	NUP214	214	P35658	K1037 [PFAK(me1)SHL]
pan-Kme1	NUP214	214	P35658	K640 [VPLK(me1)SSV]
pan-Kme1	MYLK	211	Q15746	K1101 [PAFK(me1)QKL]
pan-Kme1	MYLK	211	Q15746	K1103 [FKQK(me1)LQD]
SY46	TNRC6A	210	Q8NDV7-3	K5 [LQDK(me1)RME]
pan-Kme1	PDCD11	209	Q14690	K219 [PLLK(me1)AQE]
pan-Kme1	MLLT4	207	P55196-1	K1699 [SYLK(me1)TQV]

Tab. 7.10: Detected lysine mono-methylation sites

Antibody	Gene	Mass, kDa	Uniprot ID	Kme1 site
pan-Kme1	ECM29	204	Q5VYK3	K1099 [NSRK(me1)GAA]
pan-Kme1	ARFGEF2	202	Q9Y6D5	K1026 [GSLK(me1)GHT]
pan-Kme1	RBBP6	202	Q7Z6E9	K208 [PNMK(me1)GAM]
pan-Kme1	SIPA1L1	200	O43166	K776 [TFPK(me1)SNV]
pan-Kme1	CARD11	197	Q9BXL7	K1050 [AKNK(me1)HCL]
pan-Kme1	POLR1A	195	O95602	K1049 [VIMK(me1)SQH]
pan-Kme1	ITSN2	193	Q9NZM3	K902 [ENLK(me1)AQA]
pan-Kme1	ARAP2	193	Q8WZ64	K439 [RTQK(me1)ALI]
pan-Kme1	LMO7	193	Q8WWI1-3	K1241 [GSVK(me1)TST]
SY46	RICTOR	192	Q6R327	K1304 [QSLK(me1)APS]
pan-Kme1	ARHGEF28	192	Q8N1W1	K545 [LQSK(me1)ESL]
pan-Kme1	PIK3C2A	191	O00443	K326 [VNGK(me1)SLS]
pan-Kme1	TIAM2	190	Q8IVF5	K69 [SHFK(me1)SNQ]
pan-Kme1	IQGAP1	189	P46940	K942 [KKNK(me1)EQL]
pan-Kme1	FAM208A	189	Q9UK61	K360 [TLWK(me1)GQL]
pan-Kme1	AIM1	189	Q9Y4K1	K988 [GLFK(me1)SSR]
SY46	ADCY10	187	Q96PN6-3	K4 [MLFK(me1)AYM]
pan-Kme1	EIF2AK4	187	Q9P2K8	K535 [QLLK(me1)HSF]
pan-Kme1	KIF14	186	Q15058	K41 [LHLK(me1)SDM]
pan-Kme1	IQGAP3	185	Q86VI3	K1152 [KVLK(me1)ATL]
pan-Kme1	MADD	183	Q8WXG6-3	K884 [PSLK(me1)GNR]
pan-Kme1	TOP2B	183	Q02880	K412 [FGSK(me1)CQL]
pan-Kme1	TOP2B	183	Q02880	K957 [QVYK(me1)EQV]
pan-Kme1	DNMT1	183	P26358	K142 [RRSK(me1)SDG]
both	TNKS1BP1	182	Q9C0C2	K1641 [LGTK(me1)GLK]
both	TNKS1BP1	182	Q9C0C2	K1644 [KGLK(me1)VNL]
both	API5	182	Q9BZZ5-2	K487 [PSGK(me1)YSS]
pan-Kme1	TNKS1BP1	182	Q9C0C2	K80 [SARK(me1)MNM]
pan-Kme1	NHS	179	Q6T4R5	K1591 [FSSK(me1)SFA]
pan-Kme1	CDC42EP4	179	Q9H3Q1	K5 [PILK(me1)QLV]
pan-Kme1	CAMSAP1	178	Q5T5Y3	K211 [VKLK(me1)QQL]
pan-Kme1	CAMSAP1	178	Q5T5Y3	K912 [KLGK(me1)AAF]
SY46	DNMBP	177	Q6XZF7	K1009 [SHLK(me1)HLT]
pan-Kme1	EIF4G3	177	O43432	K1077 [SWGK(me1)GSS]
pan-Kme1	TOP2A	174	P11388	K936 [QTYK(me1)EQV]
SY46	BAI1	174	O14514	K1126 [ITDK(me1)KLK]
pan-Kme1	EPRS	171	P07814	K861 [LSLK(me1)AQY]
pan-Kme1	TULP4	169	Q9NRJ4	K1014 [PLAK(me1)SKG]
pan-Kme1	TULP4	169	Q9NRJ4	K997 [ATLK(me1)MAQ]
pan-Kme1	ERCC6	168	Q03468	K729 [TAYK(me1)CAC]
pan-Kme1	CAMSAP2	168	Q08AD1	K368 [NAAK(me1)RNV]
both	GLI2	168	P10070	K738 [EKLK(me1)SLK]
both	GLI2	168	P10070	K741 [KSLK(me1)DSC]
pan-Kme1	CLIP1	162	P30622	K509 [VSEK(me1)SRI]
pan-Kme1	CLIP1	162	P30622	K517 [ELEK(me1)DLA]
pan-Kme1	ROCK2	161	O75116	K1368 [TSMK(me1)IQQ]

Tab. 7.10: Detected lysine mono-methylation sites

Antibody	Gene	Mass, kDa	Uniprot ID	Kme1 site
pan-Kme1	FGD6	161	Q6ZV73	K175 [VVLK(me1)ASV]
pan-Kme1	FGD6	161	Q6ZV73	K650 [FWSK(me1)SSQ]
pan-Kme1	DAPK1	160	P53355	K700 [GSGK(me1)TTL]
pan-Kme1	DAPK1	160	P53355	K708 [ESLK(me1)CGL]
pan-Kme1	BLM	159	P54132	K195 [NFFK(me1)AQL]
pan-Kme1	ERBB2IP	158	Q96RT1	K869 [SVVK(me1)SHS]
pan-Kme1	ERBB2IP	158	Q96RT1	K914 [VRSK(me1)SAT]
pan-Kme1	BCL9L	157	Q86UU0	K83 [HGAK(me1)ANQ]
pan-Kme1	INSR	156	P06213	K1352 [LGFK(me1)RSY]
pan-Kme1	TAOK3	156	Q9H2K8	K738 [KALK(me1)NHQ]
pan-Kme1	ANKRD50	156	Q9ULJ7	K1185 [NSLK(me1)SSK]
pan-Kme1	POGZ	155	Q7Z3K3	K480 [DGGK(me1)VAQ]
pan-Kme1	NUP153	154	P49790	K250 [SILK(me1)TSQ]
pan-Kme1	NUP153	154	P49790	K292 [RQMK(me1)AKQ]
pan-Kme1	NUP153	154	P49790	K294 [MKAK(me1)QLS]
pan-Kme1	RAD50	154	Q92878	K458 [NELK(me1)NVK]
pan-Kme1	TECPR2	154	O15040	K1101 [HVAK(me1)GSL]
pan-Kme1	MSH6	153	P52701	K1092 [LELK(me1)GSR]
both	WHSC1	152	O96028	K582 [SSLK(me1)SQA]
SY46	WHSC1	152	O96028	K588 [AATK(me1)NLS]
pan-Kme1	ACIN1	152	Q9UKV3	K81 [TDLK(me1)AAL]
pan-Kme1	PARD3	151	Q8TEW0	K1018 [GMLK(me1)GLG]
pan-Kme1	PARD3	151	Q8TEW0-10	K984 [GMLK(me1)GLG]
pan-Kme1	KIF7	151	Q2M1P5	K869 [LELK(me1)HEQ]
pan-Kme1	PLCG1	149	P19174	K178 [KDLK(me1)NML]
SY46	RNF123	149	Q5XPI4	K991 [HLLK(me1)TKL]
SY46	RNF123	149	Q5XPI4	K993 [LKTK(me1)LED]
pan-Kme1	CLUH	147	O75153	K384 [AIFK(me1)VHS]
SY46	SPAG9	146	O60271	K577 [VNLK(me1)YNA]
pan-Kme1	VPS33B	143	Q9H267	K419 [RSLK(me1)TQY]
pan-Kme1	TRAPPC10	142	P48553	K507 [GALK(me1)NYL]
SY46	PHLDB2	142	Q86SQ0	K50 [LRFK(me1)ANG]
pan-Kme1	DHX9	141	Q08211	K1024 [LIHK(me1)SSV]
pan-Kme1	MTR	141	Q99707	K461 [AGLK(me1)CCQ]
pan-Kme1	EIF5B	139	O60841	K424 [ATLK(me1)LLQ]
pan-Kme1	TIMELESS	139	Q9UNS1	K288 [QGLK(me1)SIG]
pan-Kme1	EDRF1	139	Q3B7T1	K140 [PYSK(me1)SHV]
pan-Kme1	SSFA2	138	P28290	K701 [AQMK(me1)VCS]
pan-Kme1	TPP2	138	P29144	K1187 [FAYK(me1)HAL]
pan-Kme1	TPP2	138	P29144	K467 [SGLK(me1)ANN]
pan-Kme1	SKIV2L	138	Q15477	K837 [NGLK(me1)SLS]
SY46	RAB11FIP1	137	Q6WKZ4	K284 [TDLK(me1)QLN]
both	RAPH1	135	Q70E73	K129 [HTLK(me1)HGT]
both	RAPH1	135	Q70E73	K134 [GTLK(me1)GLS]
pan-Kme1	VCPIP1	134	Q96JH7	K1028 [FQGK(me1)GHS]
pan-Kme1	POLR2B	134	P30876	K445 [DGLK(me1)YSL]

Tab. 7.10: Detected lysine mono-methylation sites

Antibody	Gene	Mass, kDa	Uniprot ID	Kme1 site
pan-Kme1	RTEL1	134	Q9NZ71	K216 [RNLK(me1)QQA]
SY46	NAA15	133	Q9BXJ9	K756 [TFLK(me1)RNS]
pan-Kme1	RECQL4	133	O94761	K110 [QRLK(me1)ANL]
pan-Kme1	RECQL4	133	O94761	K114 [ANLK(me1)GTL]
both	FNDC3B	133	Q53EP0	K215 [SIYK(me1)SSC]
SY46	FNDC3B	133	Q53EP0	K227 [GYGK(me1)GHS]
pan-Kme1	REXO1	132	Q8N1G1	K921 [EDLK(me1)GAA]
pan-Kme1	ARHGEF18	131	Q6ZSZ5	K1080 [FDLK(me1)QQL]
pan-Kme1	BMP2K	129	Q9NSY1-2	K656 [GHLK(me1)AYF]
pan-Kme1	ENAM	129	Q9NRM1	K21 [LVPK(me1)GKM]
pan-Kme1	ENAM	129	Q9NRM1	K23 [PKGK(me1)MKI]
SY46	TBC1D8B	129	Q0IIM8	K453 [KMLK(me1)EKM]
pan-Kme1	RBM6	129	P78332	K36 [PPLK(me1)SHA]
SY46	RASAL2	129	Q9UJF2	K662 [HKLK(me1)SPS]
pan-Kme1	INPP5F	128	Q9Y2H2	K774 [AQGK(me1)NFL]
pan-Kme1	RFC1	128	P35251	K645 [SSFK(me1)AAL]
both	MICAL2	127	O94851-4	K646 [SLAK(me1)SSI]
pan-Kme1	RGSL1	126	A5PLK6	K616 [EIIK(me1)ETK]
pan-Kme1	TEX2	125	Q8IWB9	K383 [IELK(me1)SSQ]
pan-Kme1	POLD1	124	P28340	K414 [QTLK(me1)VQT]
pan-Kme1	RAB3GAP2	124	Q9H2M9	K913 [QTSK(me1)VSS]
pan-Kme1	NEMF	123	O60524	K311 [IDLK(me1)ALQ]
pan-Kme1	AASDH	123	Q4L235	K720 [QNLK(me1)GLN]
pan-Kme1	USP28	122	Q96RU2	K402 [YRSK(me1)ELI]
both	LIMCH1	122	Q9UPQ0	K588 [SLTK(me1)SQM]
pan-Kme1	SECISBP2L	122	Q93073	K176 [VVPK(me1)QQL]
both	RABGAP1	122	Q9Y3P9	K445 [PFSK(me1)RST]
pan-Kme1	MAP4	121	P27816	K933 [PDLK(me1)NVR]
pan-Kme1	USP53	121	Q70EK8	K536 [SSSK(me1)SQI]
SY46	FAM120C	121	Q9NX05	K674 [TQRK(me1)MER]
pan-Kme1	AP3B2	119	Q13367	K120 [RGLK(me1)DPN]
pan-Kme1	U2SURP	118	O15042	K392 [ERLK(me1)NPN]
pan-Kme1	U2SURP	118	O15042	K82 [NKLK(me1)AFS]
pan-Kme1	PHF8	118	Q9UPP1	K486 [NVGK(me1)TSN]
pan-Kme1	UBA1	118	P22314	K304 [ISFK(me1)SLV]
pan-Kme1	ANKRD27	117	Q96NW4	K603 [TPLK(me1)CAL]
pan-Kme1	MPRIP	117	Q6WCQ1	K200 [PTTK(me1)STL]
pan-Kme1	PSD3	116	Q9NYI0	K124 [SHLK(me1)EQS]
pan-Kme1	BMPR2	115	Q13873	K219 [AVYK(me1)GSL]
pan-Kme1	ANKRD52	115	Q8NB46	K1036 [SLLK(me1)NCS]
pan-Kme1	NNT	114	Q13423	K6 [NLLK(me1)TVV]
SY46	NEFH	112	P12036	K663 [SPEK(me1)EEA]
pan-Kme1	USP15	112	Q9Y4E8	K259 [RNVK(me1)NSN]
pan-Kme1	ZC3H7B	112	Q9UGR2	K689 [HAGK(me1)ASS]
pan-Kme1	ARHGEF2	112	Q92974	K100 [ALLK(me1)NNT]
pan-Kme1	SEC23IP	111	Q9Y6Y8	K871 [SDLK(me1)QGF]

Tab. 7.10: Detected lysine mono-methylation sites

Antibody	Gene	Mass, kDa	Uniprot ID	Kme1 site
pan-Kme1	DENND1A	111	Q8TEH3	K314 [AFLK(me1)AQA]
pan-Kme1	ZBED6	110	P86452	K250 [FLIK(me1)SNI]
pan-Kme1	INPP4A	110	Q96PE3	K321 [PSFK(me1)ASS]
pan-Kme1	SMG8	110	Q8ND04	K242 [PLLK(me1)TAI]
pan-Kme1	DIS3	109	Q9Y2L1	K331 [RMLK(me1)TAV]
pan-Kme1	DIS3	109	Q9Y2L1	K585 [INSK(me1)ASL]
pan-Kme1	COG1	109	Q8WTW3	K506 [LSMK(me1)AQA]
both	THRAP3	109	Q9Y2W1	K252 [PALK(me1)SPL]
pan-Kme1	RBM15	107	Q96T37	K420 [GFLK(me1)FEN]
pan-Kme1	ZC3H18	106	Q86VM9	K229 [FFMK(me1)GNC]
pan-Kme1	SLK	105	Q9H2G2	K379 [LNSK(me1)ILN]
pan-Kme1	BBX	105	Q8WY36	K518 [SSGK(me1)GSI]
pan-Kme1	CASK	105	O14936	K305 [RKLK(me1)GAV]
pan-Kme1	CASK	105	O14936	K703 [WFGK(me1)KKK]
pan-Kme1	ZNF618	105	Q5T7W0	K950 [MFLK(me1)SNM]
pan-Kme1	PACS1	105	Q6VY07	K903 [VDSK(me1)SQV]
pan-Kme1	ZFPM1	105	Q8IX07	K168 [IHRK(me1)DDA]
pan-Kme1	CUL4B	104	Q13620-1	K142 [LKNK(me1)SIL]
pan-Kme1	GPATCH1	103	Q9BRR8	K456 [TDLK(me1)AAQ]
pan-Kme1	TUBGCP2	103	Q9BSJ2	K714 [KNLK(me1)SAS]
both	EIF4G2	102	P78344	K431 [KFMK(me1)SQG]
pan-Kme1	EIF4G2	102	P78344	K464 [FSKK(me1)GQL]
pan-Kme1	KIF18A	102	Q8NI77	K360 [SSLK(me1)SNV]
pan-Kme1	TRAK2	101	O60296	K831 [GQLK(me1)MNL]
SY46	WDR44	101	Q5JSH3	K486 [VKFK(me1)AAH]
SY46	WDR44	101	Q5JSH3	K492 [HGFK(me1)GPY]
pan-Kme1	KIFC1	101	Q9BW19	K308 [QELK(me1)GNI]
pan-Kme1	ADNP	101	Q9H2P0	K408 [GQLK(me1)SPS]
pan-Kme1	GTF3C2	101	Q8WUA4	K148 [RKSK(me1)AEL]
pan-Kme1	CORO7	101	P57737	K491 [TNLK(me1)GLN]
pan-Kme1	DDHD1	100	Q8NEL9	K744 [NIGK(me1)ASI]
SY46	SYNPO	99	Q8N3V7	K599 [SHLK(me1)GQA]
SY46	SYNPO	99	Q8N3V7	K607 [PASK(me1)TGI]
pan-Kme1	SMTN	99	P53814	K751 [ELMK(me1)AQS]
pan-Kme1	PPP1R10	99	Q96QC0	K918 [FMMK(me1)GNC]
pan-Kme1	LRRC8D	98	Q7L1W4	K806 [LELK(me1)GNC]
pan-Kme1	DNM2	98	P50570	K44 [SAGK(me1)SSV]
pan-Kme1	COPG1	98	Q9Y678	K145 [RYMK(me1)QAI]
pan-Kme1	COPG2	98	Q9UBF2	K145 [RYMK(me1)QAI]
pan-Kme1	SMEK2	97	Q5MIZ7	K753 [ETKK(me1)AKE]
pan-Kme1	SMEK2	97	Q5MIZ7	K755 [KKAK(me1)ESE]
SY46	AKAP1	97	Q92667	K440 [SCLK(me1)SLL]
pan-Kme1	RHBDF2	97	Q6PJF5	K87 [AYLK(me1)SVS]
pan-Kme1	EXD3	97	Q8N9H8	K702 [CSLK(me1)AQQ]
pan-Kme1	PRICKLE2	96	Q7Z3G6	K820 [GLPK(me1)SST]
pan-Kme1	PAN3	96	Q58A45	K496 [RIQK(me1)SSN]

Tab. 7.10: Detected lysine mono-methylation sites

Antibody	Gene	Mass, kDa	Uniprot ID	Kme1 site
pan-Kme1	SMEK1	95	Q6IN85	K733 [VLLK(me1)TNL]
both	ILF3	95	Q12906	K613 [GGPK(me1)FAA]
pan-Kme1	HSPA4L	95	O95757	K53 [NAAK(me1)SQI]
pan-Kme1	HSPA4	94	P34932	K53 [AAAK(me1)SQV]
pan-Kme1	CC2D1B	94	Q5T0F9	K785 [IFHK(me1)GSF]
pan-Kme1	CCSER2	94	Q9H7U1	K115 [GGLK(me1)SVS]
pan-Kme1	DOCK5	94	Q9H7D0	K1741 [SLSK(me1)SQV]
SY46	CALD1	93	Q05682	K504 [HKLK(me1)HTE]
both	BCAR3	93	O75815	K334 [LSLK(me1)AHQ]
pan-Kme1	HSP90B1	92	P14625	K663 [RIMK(me1)AQA]
pan-Kme1	RBM5	92	P52756	K780 [VRLK(me1)GAG]
pan-Kme1	RBM5	92	P52756	K787 [LGAK(me1)GSA]
pan-Kme1	ANAPC4	92	Q9UJX5	K718 [ESMK(me1)AQY]
pan-Kme1	AFAP1L2	91	Q8N4X5	K332 [AGLK(me1)LSN]
SY46	BANK1	89	Q8NDB2	K79 [TSYK(me1)CKL]
pan-Kme1	PPP1R13L	89	Q8WUF5	K28 [MDLK(me1)QME]
SY46	PPP1R13L	89	Q8WUF5	K522 [RPLK(me1)RRG]
pan-Kme1	VEZT	89	Q9HBM0	K593 [LGFK(me1)ASE]
pan-Kme1	AFG3L2	89	Q9Y4W6	K618 [LYTK(me1)EQL]
pan-Kme1	MAPK7	88	Q13164	K480 [AALK(me1)AAL]
both	IFI16	88	Q16666	K186 [ENPK(me1)TVA]
pan-Kme1	AARSD1	88	Q9BTE6-2	K378 [ALLK(me1)CGA]
pan-Kme1	DIEXF	87	Q68CQ4	K245 [IDLK(me1)SLH]
pan-Kme1	ERCC2	87	P18074	K133 [VDGK(me1)CHS]
pan-Kme1	TLK1	87	Q9UKI8-3	K3 [_MLK(me1)LAA]
pan-Kme1	AFAP1L1	86	Q8TED9	K383 [AELK(me1)GSM]
both	TBC1D12	86	O60347	K335 [KELK(me1)SVV]
both	TBC1D12	86	O60347	K345 [PGWK(me1)LFG]
SY46	ACO2	85	Q99798	K409 [HGLK(me1)CKS]
pan-Kme1	EZH2	85	Q15910	K735 [DALK(me1)YVG]
pan-Kme1	LIMA1	85	Q9UHB6	K292 [NELK(me1)ASG]
SY46	LIMA1	85	Q9UHB6	K447 [QLFK(me1)SKG]
SY46	LIMA1	85	Q9UHB6	K603 [TSVK(me1)SPK]
pan-Kme1	KIF3B	85	O15066	K583 [LKLK(me1)HLI]
pan-Kme1	TATDN2	85	Q93075	K189 [IYLK(me1)AIQ]
both	HSP90AA1	85	P07900	K615 [RIMK(me1)AQA]
pan-Kme1	FAM129B	84	Q96TA1	K401 [HPLK(me1)MQS]
SY46	MAP7	84	Q14244	K136 [EEDK(me1)ERH]
pan-Kme1	GYS1	84	P13807	K381 [ETLK(me1)GQA]
pan-Kme1	BOP1	84	Q14137	K711 [KVLK(me1)GHV]
both	HSP90AB1	83	P08238	K607 [RIMK(me1)AQA]
pan-Kme1	THNSL1	83	Q8IYQ7	K568 [DILK(me1)SSN]
pan-Kme1	GOLGA5	83	Q8TBA6	K74 [RNQK(me1)ATI]
pan-Kme1	OAS2	82	P29728	K419 [NSLK(me1)SYT]
pan-Kme1	MCM5	82	P33992	K141 [RSLK(me1)SDM]
pan-Kme1	C1orf168	82	Q5VWT5	K604 [DKLK(me1)MWK]

Tab. 7.10: Detected lysine mono-methylation sites

Antibody	Gene	Mass, kDa	Uniprot ID	Kme1 site
pan-Kme1	EPB41L5	82	Q9HCM4	K563 [PDFK(me1)SNI]
pan-Kme1	EPB41L5	82	Q9HCM4	K568 [NILK(me1)AQV]
SY46	SLC26A2	82	P50443	K503 [DLPK(me1)MWS]
pan-Kme1	KIF2C	81	Q99661	K354 [GSGK(me1)THT]
pan-Kme1	KIF2C	81	Q99661	K365 [LSGK(me1)AQN]
pan-Kme1	DDHD2	81	O94830	K587 [MDLK(me1)NNL]
pan-Kme1	AKAP17A	81	Q02040	K56 [ERLK(me1)GMV]
pan-Kme1	RTF1	80	Q92541	K587 [HNMK(me1)NQQ]
pan-Kme1	DTL	79	Q9NZJ0	K537 [VSQK(me1)SSQ]
pan-Kme1	BBS12	79	Q6ZW61	K217 [RTLK(me1)NSL]
pan-Kme1	MPHOSPH10	79	O00566	K641 [KALK(me1)SSQ]
pan-Kme1	UVRAG	78	Q9P2Y5	K455 [PNLK(me1)NFM]
pan-Kme1	PRPF3	78	O43395	K282 [PTLK(me1)ANI]
both	CAST	77	P20810-9	K303 [EVLK(me1)AQS]
pan-Kme1	NUFIP2	76	Q7Z417	K136 [TSLK(me1)QTV]
pan-Kme1	TMPO	75	P42166	K674 [TPFK(me1)GGT]
pan-Kme1	ZNF800	75	Q2TB10	K58 [KQLK(me1)HIL]
both	GAS2L3	75	Q86XJ1	K360 [SSLK(me1)GGN]
pan-Kme1	FIGNL1	74	Q6PIW4	K126 [KKFK(me1)DSL]
pan-Kme1	FIGNL1	74	Q6PIW4	K296 [PTFK(me1)TAK]
pan-Kme1	FIGNL1	74	Q6PIW4	K299 [KTAK(me1)EQL]
pan-Kme1	PATZ1	74	Q9HBE1	K272 [RPRK(me1)ANL]
pan-Kme1	SARG	74	Q9BW04	K511 [SLGK(me1)GSF]
pan-Kme1	SLC24A2	74	Q9UI40	K288 [QVEK(me1)WVK]
pan-Kme1	SLC24A2	74	Q9UI40	K291 [KWVK(me1)QMI]
pan-Kme1	DDX3X	73	O00571	K81 [SRGK(me1)SSF]
pan-Kme1	AGPS	73	O00116	K628 [QWLK(me1)ESI]
pan-Kme1	TNK1	72	Q13470	K475 [DRKK(me1)ANL]
both	HSPA5	72	P11021	K585 [LGGK(me1)LSS]
SY46	SAMHD1	72	Q9Y3Z3	K312 [DVDK(me1)WDY]
pan-Kme1	THEMIS2	72	Q5TEJ8	K573 [GGVK(me1)SSQ]
pan-Kme1	MTMR6	72	Q9Y217	K565 [PNLK(me1)TSL]
pan-Kme1	CD2AP	71	Q9Y5K6	K397 [PPTK(me1)ASN]
both	CD2AP	71	Q9Y5K6	K557 [SASK(me1)ANT]
pan-Kme1	TRIM25	71	Q14258	K416 [VDLK(me1)QAG]
SY46	SCG2	71	P13521	K274 [NIEK(me1)NEQ]
SY46	SCG2	71	P13521	K283 [DEMK(me1)RSG]
both	TXNRD1	71	Q16881	K138 [PTLK(me1)AYQ]
SY46	HSPA8	71	P11142	K251 [KHKK(me1)DIS]
pan-Kme1	HSPA8	71	P11142	K3 [_MSK(me1)GPA]
pan-Kme1	PABPC4	71	Q13310	K375 [EERK(me1)AHL]
pan-Kme1	LBR	71	Q14739	K147 [APHK(me1)NTQ]
pan-Kme1	WNT5B	71	Q9H1J7	K327 [NQFK(me1)SVQ]
pan-Kme1	VPS33B	71	Q9H267	K59 [SILK(me1)QHE]
pan-Kme1	HSPA1A	70	P0DMV8	K3 [_MAK(me1)AAA]
pan-Kme1	PHACTR2	70	O75167	K29 [PPFK(me1)RKG]

Tab. 7.10: Detected lysine mono-methylation sites

Antibody	Gene	Mass, kDa	Uniprot ID	Kme1 site
both	DENND6A	70	Q8IWF6	K510 [HFLK(me1)SPN]
pan-Kme1	TRIM47	70	Q96LD4	K361 [SFTK(me1)SSQ]
pan-Kme1	BTBD9	69	Q96Q07-2	K344 [PELK(me1)QSS]
pan-Kme1	DDX5	69	P17844	K197 [CRLK(me1)STC]
pan-Kme1	FAM107B	69	Q9H098	K25 [NPVK(me1)TSR]
pan-Kme1	ARHGAP40	69	Q5TG30	K145 [SGMK(me1)GAQ]
pan-Kme1	ZNF143	69	P52747	K285 [YGLK(me1)SHV]
SY46	SNX18	69	Q96RF0	K290 [TKFK(me1)GMK]
both	SNX18	69	Q96RF0	K293 [KGMK(me1)SYI]
pan-Kme1	GNL1	69	P36915	K373 [NVGK(me1)SSL]
SY46	EPN2	68	O95208	K423 [SAGK(me1)RAS]
pan-Kme1	PPP1R18	68	Q6NYC8	K8 [PDWK(me1)LQL]
pan-Kme1	ZNF668	68	Q96K58	K337 [RPFK(me1)CLQ]
pan-Kme1	MSN	68	P26038	K143 [EVHK(me1)SGY]
pan-Kme1	STXBP3	68	O00186	K11 [RGLK(me1)SVV]
pan-Kme1	ARAF	68	P10398	K431 [RDLK(me1)SNN]
pan-Kme1	WDR1	66	O75083	K334 [NGGK(me1)SYI]
pan-Kme1	NOP56	66	O00567	K375 [LANK(me1)CSI]
pan-Kme1	KLC1	65	Q07866-6	K548 [GSLK(me1)RSG]
pan-Kme1	CDKAL1	65	Q5VV42	K150 [DYLK(me1)GLS]
pan-Kme1	FRMD5	65	Q7Z6J6	K313 [LFFK(me1)GSR]
pan-Kme1	CHMP4B	65	Q9H444	K107 [EVLK(me1)NMG]
pan-Kme1	BIN1	65	O00499	K7 [MGSK(me1)GVT]
pan-Kme1	C3orf17	65	Q6NW34	K332 [FDFK(me1)CSQ]
pan-Kme1	ASNS	64	P08243	K467 [WRPK(me1)EAF]
pan-Kme1	ASNS	64	P08243	K556 [THYK(me1)SAV]
pan-Kme1	SARG	64	Q9BW04	K480 [MNFK(me1)SNT]
pan-Kme1	RIOK2	64	Q9BVS4	K531 [IFTK(me1)QRR]
pan-Kme1	PDLIM5	64	Q96HC4	K116 [AVSK(me1)VTS]
pan-Kme1	RAVER1	64	Q8IY67	K587 [PAQK(me1)AAM]
pan-Kme1	RAVER1	64	Q8IY67-2	K657 [AGLK(me1)QSH]
pan-Kme1	ZFP91	63	Q96JP5	K174 [SRSK(me1)TGS]
pan-Kme1	UTP14A	63	Q9BVJ6	K136 [AFNK(me1)TAQ]
pan-Kme1	RIOK2	63	Q9BVS4	K541 [QNIK(me1)SSL]
pan-Kme1	CPNE8	63	Q86YQ8	K405 [RSLK(me1)SVQ]
pan-Kme1	NARS	63	O43776	K31 [TGLK(me1)ALM]
pan-Kme1	WDR20	63	Q8TBZ3	K446 [AGSK(me1)SSV]
pan-Kme1	SLC22A5	63	O76082	K553 [TILK(me1)STA]
both	PIP5K1A	63	Q99755	K86 [SALK(me1)GAI]
both	CAMK2G	63	Q13555	K301 [RKLK(me1)GAI]
pan-Kme1	GMEB1	63	Q9Y692	K359 [HRLK(me1)SQT]
pan-Kme1	SLAIN2	63	Q9P270	K182 [SALK(me1)RQN]
pan-Kme1	SLAIN2	63	Q9P270	K429 [DSVK(me1)SSR]
pan-Kme1	GLUD1	61	P00367	K171 [MTYK(me1)CAV]
pan-Kme1	FAM114A1	61	Q8IWE2	K135 [SWGK(me1)SLL]
pan-Kme1	PAF1	60	Q8N7H5	K62 [VQYK(me1)ATS]

Tab. 7.10: Detected lysine mono-methylation sites

Antibody	Gene	Mass, kDa	Uniprot ID	Kme1 site
pan-Kme1	CCT7	59	Q99832	K430 [IPGK(me1)QQL]
pan-Kme1	OASL	59	Q15646	K331 [QCLK(me1)QDC]
both	ATE1	59	O95260	K285 [SQFK(me1)ATL]
pan-Kme1	KIAA1671	59	Q9BY89	K853 [RAIK(me1)AAI]
pan-Kme1	SPATA2	58	Q9UM82	K510 [LNYK(me1)STQ]
pan-Kme1	CCT4	58	P50991	K65 [QDGK(me1)GDV]
both	ZNF513	58	Q8N8E2	K166 [SHLK(me1)RHM]
pan-Kme1	DKC1	58	O60832	K394 [LMIK(me1)QGL]
pan-Kme1	ARFGAP2	57	Q8N6H7	K209 [LELK(me1)SSI]
pan-Kme1	DYNC1LI1	57	Q9Y6G9	K435 [PNMK(me1)AGA]
SY46	CAMK2D	56	Q13557	301 [RKLK(me1)GAI]
pan-Kme1	CAMK2D	56	Q13557	K301 [RKLK(me1)GAI]
pan-Kme1	LAP3	56	P28838-2	K3 [_MTK(me1)GLV]
pan-Kme1	CBX2	56	Q14781	K277 [LALK(me1)AQA]
pan-Kme1	SHMT2	56	P34897	K469 [QDFK(me1)SFL]
pan-Kme1	IMPDH2	56	P12268	K195 [ITLK(me1)EAN]
pan-Kme1	SRP54	56	P61011	K496 [GNMK(me1)GMM]
pan-Kme1	PRPF31	55	Q8WWY3	K438 [YGGK(me1)STI]
pan-Kme1	RTCB	55	Q9Y3I0	K308 [DYLK(me1)GMA]
pan-Kme1	PRPF19	55	Q9UMS4	K428 [KNFK(me1)TLQ]
pan-Kme1	ZMPSTE24	55	O75844	K437 [ALIK(me1)LNK]
pan-Kme1	ICA1	55	Q05084	K118 [ATGK(me1)ALC]
pan-Kme1	YAP1	54	P46937	K321 [LRLK(me1)QQE]
SY46	WASF2	54	Q9Y6W5	K109 [QDQK(me1)LFD]
pan-Kme1	NONO	54	Q15233	K5 [QSNK(me1)TFN]
pan-Kme1	DLD	54	P09622	K66 [LGFK(me1)TVC]
pan-Kme1	FBXW9	54	Q5XUX1-3	K301 [ALLK(me1)HQQ]
pan-Kme1	ZBTB45	54	Q96K62	K448 [YLLK(me1)HMV]
pan-Kme1	ZNF518B	54	Q9C0D4	K482 [VALK(me1)GHS]
pan-Kme1	ETV4	54	P43268	K6 [RRMK(me1)AGY]
pan-Kme1	TRIM38	53	O00635	K222 [LGLK(me1)SNE]
pan-Kme1	ACOT2	53	P49753	K39 [WSLK(me1)SSA]
pan-Kme1	EPHX1	53	P07099	K117 [PHFK(me1)TKI]
pan-Kme1	EPHX1	53	P07099	K119 [FKTK(me1)IEG]
pan-Kme1	MINA	53	Q8IUF8	K87 [TDLK(me1)SLC]
pan-Kme1	PPHLN1	53	Q8NEY8-2	K206 [QSLK(me1)TSR]
pan-Kme1	SNX8	53	Q9Y5X2	K103 [QRFK(me1)SSV]
pan-Kme1	DGCR14	53	Q96DF8	K358 [LGLK(me1)MAN]
SY46	DGCR14	53	Q96DF8	K429 [THLK(me1)TPA]
both	SAMM50	52	Q9Y512	K255 [HSLK(me1)SSL]
pan-Kme1	SNX4	52	O95219	K356 [FSLK(me1)GMT]
both	CAP1	52	Q01518	K287 [PALK(me1)AQS]
pan-Kme1	NSUN6	52	Q8TEA1	K192 [PELK(me1)GMG]
pan-Kme1	POLD3	51	Q15054	K231 [APGK(me1)GNM]
pan-Kme1	ODR4	51	Q5SWX8	K282 [VNLK(me1)GAV]
SY46	TMPO	51	P42167	K303 [GNFK(me1)HAS]

Tab. 7.10: Detected lysine mono-methylation sites

Antibody	Gene	Mass, kDa	Uniprot ID	Kme1 site
pan-Kme1	ERRFI1	51	Q9UJM3	K16 [VPLK(me1)TGF]
pan-Kme1	ERRFI1	51	Q9UJM3	K41 [SEFK(me1)NNF]
pan-Kme1	EEF1A2	50	Q05639	K165 [YSEK(me1)RYD]
pan-Kme1	EEF1A2	50	Q05639	K318 [VSVK(me1)DIR]
both	EEF1A2	50	Q05639	K79 [SLWK(me1)FET]
both	EEF1A2	50	Q05639	K84 [ETTK(me1)YYI]
SY46	SH3BP5	50	O60239	K217 [LKAK(me1)YYV]
both	EEF1A1	50	P68104	K165 [YSQK(me1)RYE]
both	EEF1A1	50	P68104	K172 [EIVK(me1)EVS]
pan-Kme1	EEF1A1	50	P68104	K318 [VSVK(me1)DVR]
both	EEF1A1	50	P68104	K55 [GSFK(me1)YAW]
pan-Kme1	TUBAL3	50	A6NHL2	K361 [LDHK(me1)FDL]
pan-Kme1	TBL2	50	Q9Y4P3	K276 [FELK(me1)GHS]
pan-Kme1	SDE2	50	Q6IQ49	K173 [SVLK(me1)GMQ]
pan-Kme1	HNRNPH1	49	P31943	K185 [EIFK(me1)SSR]
pan-Kme1	ARMCX1	49	Q9P291	K116 [NTLK(me1)EQA]
pan-Kme1	VPS4A	49	Q9UN37	K23 [DKAK(me1)NYE]
pan-Kme1	ACP6	49	Q9NPH0	K116 [TTLK(me1)GGM]
pan-Kme1	LRRC42	49	Q9Y546	K349 [APLK(me1)CPL]
pan-Kme1	KHDRBS1	48	Q07666	K432 [RPVK(me1)GAY]
both	DBNL	48	Q9UJU6	K164 [VYQK(me1)TNA]
pan-Kme1	PIP4K2C	47	Q8TBX8	K220 [YDLK(me1)GSL]
pan-Kme1	ENO1	47	P06733	K343 [LLLK(me1)VNQ]
pan-Kme1	ENO1	47	P06733	K406 [RLAK(me1)YNQ]
SY46	SENP3	47	Q9H4L4	K43 [PRLK(me1)SGG]
pan-Kme1	PLIN3	47	O60664	K166 [DKTK(me1)SVV]
pan-Kme1	EIF4A3	47	P38919	K198 [KGFK(me1)EQI]
SY46	NSUN5	47	Q96P11-2	K144 [NTLK(me1)TCS]
pan-Kme1	TBC1D13	47	Q9NVG8	K168 [TTLK(me1)SQT]
pan-Kme1	LUC7L2	47	Q9Y383	K37 [RVCK(me1)SHL]
pan-Kme1	SUCLG2	47	Q96I99	K218 [GPLK(me1)SQA]
pan-Kme1	SUCLG2	47	Q96I99	K83 [IVLK(me1)AQI]
pan-Kme1	ARRB2	46	P32121	K397 [LRLK(me1)GMK]
pan-Kme1	TSEN34	45	Q9BSV6	K87 [TSFK(me1)RQQ]
both	HSDL2	45	Q6YN16	K163 [VWFK(me1)QHC]
both	KIN	45	O60870	K221 [TSSK(me1)SST]
SY46	KIN	45	O60870	K231 [SALK(me1)TIG]
pan-Kme1	CCDC85C	45	A6NKD9	K416 [NQFK(me1)GPL]
pan-Kme1	TSPYL5	45	Q86VY4	K108 [GPGK(me1)AAS]
pan-Kme1	RBM17	45	Q96I25	K24 [KNFK(me1)LLQ]
pan-Kme1	RBM17	45	Q96I25	K383 [RVVK(me1)ACF]
pan-Kme1	OLA1	45	Q9NTK5	K24 [TSLK(me1)IGI]
pan-Kme1	YY1	45	P25490	K401 [TNLK(me1)SHI]
pan-Kme1	PGK1	45	P00558-2	K2 [__MK(me1)NNQ]
pan-Kme1	PGK1	45	P00558	K30 [VPMK(me1)NNQ]
both	MAP2K4	44	P45985	K45 [KALK(me1)LNF]

Tab. 7.10: Detected lysine mono-methylation sites

Antibody	Gene	Mass, kDa	Uniprot ID	Kme1 site
both	MAP2K4	44	P45985	K54 [PPFK(me1)STA]
SY46	RRAGC	44	Q9HB90	K387 [SSLK(me1)ALT]
pan-Kme1	PA2G4	44	Q9UQ80	K191 [HQLK(me1)QHV]
pan-Kme1	MAT2A	44	P31153	K367 [LDLK(me1)KPI]
pan-Kme1	PTRF	43	Q6NZI2	K152 [RNFK(me1)VMI]
pan-Kme1	PELO	43	Q9BRX2	K254 [YSLK(me1)EAL]
pan-Kme1	PDHA1	43	P08559	K77 [MELK(me1)ADQ]
pan-Kme1	METAP1	43	P53582	K278 [IIQK(me1)HAQ]
pan-Kme1	GTPBP10	43	A4D1E9	K161 [NAGK(me1)SSL]
pan-Kme1	UBE2Q2	43	Q8WVN8-3	K10 [TEEK(me1)LEC]
pan-Kme1	IDH3G	43	P51553	K115 [VALK(me1)GNI]
pan-Kme1	LDB2	43	O43679-2	K232 [DCLK(me1)TCL]
SY46	LDB2	43	O43679-2	K37 [EMNK(me1)RLQ]
pan-Kme1	SERPINB1	43	P30740	K213 [EDLK(me1)CRV]
SY46	DEK	43	P35659	K84 [AQGK(me1)GQK]
SY46	SAPCD2	43	Q86UD0	K233 [KQMK(me1)ELE]
pan-Kme1	KCTD9	43	Q7L273	K241 [INFK(me1)MAN]
pan-Kme1	RAD9A	43	Q99638	K2 [__MK(me1)CLV]
both	WDR74	42	Q6RFH5	K311 [VYLK(me1)SQL]
pan-Kme1	MTHFSD	42	Q2M296	K41 [PNFK(me1)GSY]
pan-Kme1	VAT1	42	Q99536	K372 [DAMK(me1)QMQ]
pan-Kme1	MBNL1	42	Q9NR56	K41 [HPSK(me1)SCQ]
pan-Kme1	AS3MT	42	Q9HBK9	K22 [QVLK(me1)RSA]
pan-Kme1	SH3GL1	41	Q99961	K12 [QFYK(me1)ASQ]
pan-Kme1	SH3GL1	41	Q99961	K67 [SRAK(me1)LTM]
pan-Kme1	SURF6	41	O75683	K7 [LLAK(me1)DAY]
pan-Kme1	AMZ2	41	Q86W34	K12 [QTLK(me1)TAL]
pan-Kme1	PEX14	41	O75381	K220 [NSLK(me1)GLL]
pan-Kme1	PSAT1	40	Q9Y617	K333 [LSLK(me1)GHR]
pan-Kme1	NUCKS1	40	Q9H1E3	K175 [PRLK(me1)ATV]
pan-Kme1	FAM50A	40	Q14320	K5 [AQYK(me1)GAA]
pan-Kme1	FAM50A	40	Q14320	K61 [AELK(me1)SST]
pan-Kme1	FAM50A	40	Q14320	K74 [NDMK(me1)AKQ]
pan-Kme1	SLC25A3	40	Q00325-2	K111 [VTLK(me1)EDG]
pan-Kme1	SLC25A3	40	Q00325-2	K233 [TMMK(me1)FAC]
pan-Kme1	DCDC1	40	P59894	K287 [IRMK(me1)KLT]
pan-Kme1	DCDC1	40	P59894	K288 [RMKK(me1)LTE]
pan-Kme1	MCUR1	40	Q96AQ8	K257 [HQLK(me1)QQV]
both	RFC4	40	P35249	K6 [AFLK(me1)GTS]
both	ALDOA	39	P04075	K322 [ENLK(me1)AAQ]
pan-Kme1	TNIP3	39	Q96KP6	K247 [SQIK(me1)ACQ]
pan-Kme1	TNIP3	39	Q96KP6	K253 [QMEK(me1)EKL]
pan-Kme1	CDC37L1	39	Q7L3B6	K62 [EFVK(me1)SSV]
pan-Kme1	AVEN	39	Q9NQS1	K230 [MQLK(me1)GPL]
pan-Kme1	AVEN	39	Q9NQS1	K245 [FELK(me1)SVA]
pan-Kme1	GPALPP1	38	Q8IXQ4	K335 [SHGK(me1)GNM]

Tab. 7.10: Detected lysine mono-methylation sites

Antibody	Gene	Mass, kDa	Uniprot ID	Kme1 site
pan-Kme1	L3HYPDH	38	Q96EM0	K299 [RAFK(me1)SSA]
pan-Kme1	TMEM245	38	Q9H330-2	K875 [EDLK(me1)SSV]
pan-Kme1	SCAMP1	38	O15126	K334 [NAFK(me1)GNQ]
SY46	PRKAG1	38	P54619	K34 [SFMK(me1)SHR]
both	CNN3	36	Q15417	K158 [GKLK(me1)AGQ]
both	CNN3	36	Q15417	K172 [GTNK(me1)CAS]
pan-Kme1	CNN3	36	Q15417	K256 [ASQK(me1)GMS]
pan-Kme1	JUN	36	P05412	K35 [KILK(me1)QSM]
pan-Kme1	SEC13	36	P55735	K203 [GQWK(me1)EEQ]
pan-Kme1	SEC13	36	P55735	K207 [EEQK(me1)LEA]
pan-Kme1	MTX3	35	Q5HYI7	K38 [APLK(me1)VNV]
pan-Kme1	GNB2L1	35	P63244	K12 [GTLK(me1)GHN]
pan-Kme1	TOMM34	35	Q15785	K241 [LVLK(me1)QYT]
pan-Kme1	NT5C3B	34	Q969T7	K10 [TLMK(me1)ATV]
pan-Kme1	NT5C3B	34	Q969T7-2	K2 [__MK(me1)ATV]
both	RPL5	34	P46777	K164 [GALK(me1)GAV]
pan-Kme1	CDK1	34	P06493	K266 [LLSK(me1)MLI]
both	CNN2	34	Q99439	K160 [ATMK(me1)AGQ]
both	CNN2	34	Q99439	K174 [GTNK(me1)CAS]
both	CFDP1	34	Q9UEE9	K219 [SGLK(me1)RSS]
both	BLVRA	33	P53004	K234 [FHFK(me1)SGS]
pan-Kme1	DTWD2	33	Q8NBA8	K271 [HLLK(me1)NGL]
pan-Kme1	ALKBH3	33	Q96Q83	K19 [APVK(me1)SQA]
pan-Kme1	CDK5	33	Q00535	K20 [TVFK(me1)AKN]
pan-Kme1	CDK5	33	Q00535	K22 [FKAK(me1)NRE]
pan-Kme1	PURB	33	Q96QR8	K136 [LDLK(me1)ENQ]
pan-Kme1	SLC25A4	33	P12235	K43 [HASK(me1)QIS]
pan-Kme1	RPSA	33	P08865	K89 [AVLK(me1)FAA]
pan-Kme1	SLC25A5	33	P05141	K166 [KIYK(me1)SDG]
pan-Kme1	SLC25A5	33	P05141	K52 [KQYK(me1)GII]
pan-Kme1	HAUS1	32	Q96CS2	K186 [DFLK(me1)AKS]
pan-Kme1	HAUS1	32	Q96CS2	K188 [LKAK(me1)SEE]
pan-Kme1	GEMIN2	32	O14893	K125 [SHWK(me1)SQQ]
pan-Kme1	CAPZB	31	P47756-2	K267 [EALK(me1)RKQ]
both	TPI1	31	P60174	K231 [GWLK(me1)SNV]
pan-Kme1	FAM122A	31	Q96E09	K203 [GPLK(me1)RKC]
pan-Kme1	CPSF4	30	O95639	K46 [FFLK(me1)AAC]
pan-Kme1	CPSF4	30	O95639	K51 [ACGK(me1)GGM]
pan-Kme1	TIGAR	30	Q9NQ88	K37 [TGFK(me1)QAA]
pan-Kme1	DNAJC9	30	Q8WXX5	K217 [DSLK(me1)AAI]
pan-Kme1	NECAP1	30	Q8NC96	K154 [LGFK(me1)EGQ]
both	LASP1	30	Q14847	K75 [LRLK(me1)QQS]
pan-Kme1	PSMA1	30	P25786	K50 [VALK(me1)RAQ]
pan-Kme1	USE1	29	Q9NZ43	K184 [RSLK(me1)TNT]
pan-Kme1	RPL7	29	P18124	K148 [PNLK(me1)SVN]
pan-Kme1	MVB12A	29	Q96EY5	K179 [QPSK(me1)GGL]

Tab. 7.10: Detected lysine mono-methylation sites

Antibody	Gene	Mass, kDa	Uniprot ID	Kme1 site
SY46	RPS6	29	P62753	K2 [_ _MK(me1)LNI]
pan-Kme1	NIPSNAP3A	28	Q9UFN0	K166 [DAFK(me1)RAV]
SY46	NECAP2	28	Q9NVZ3	K58 [AYIK(me1)LED]
SY46	VAPA	28	Q9P0L0	K26 [LKFK(me1)GPF]
pan-Kme1	PSMA7	28	O14818	K157 [HAWK(me1)ANA]
both	U2AF1	28	Q01081	K39 [LHNK(me1)PTF]
pan-Kme1	ETFB	28	P38117	K200 [NIMK(me1)AKK]
pan-Kme1	ETFB	28	P38117	K202 [MKAK(me1)KKK]
both	C12orf60	28	Q5U649	K201 [EDSK(me1)NPT]
both	C12orf60	28	Q5U649	K205 [NPTK(me1)SAA]
SY46	KLC2	27	Q9H0B6	K439 [SWYK(me1)ACK]
pan-Kme1	EFHD2	27	Q96C19	K233 [AAFK(me1)ELQ]
pan-Kme1	RABL3	26	Q5HYI8	K231 [GTLK(me1)SLH]
SY46	CHTOP	26	Q9Y3Y2	K28 [NMLK(me1)NKQ]
pan-Kme1	CHCHD3	26	Q9NX63	K203 [QTLK(me1)CSA]
pan-Kme1	QDPR	26	P09417	K109 [LMWK(me1)QSI]
both	CXorf56	26	Q9H5V9	K213 [AKMK(me1)GTL]
pan-Kme1	ALG2	26	Q9H553	K383 [PSLK(me1)ATM]
pan-Kme1	POP4	25	O95707	K2 [_ _MK(me1)SVI]
pan-Kme1	PSPH	25	P78330	K59 [VPFK(me1)AAL]
pan-Kme1	PTPN23	25	Q9H3S7	K1602 [RMSK(me1)HNF]
both	C18orf21	25	Q32NC0	K215 [SSLK(me1)GGL]
pan-Kme1	RAB11B	24	Q15907	K24 [GVGK(me1)SNL]
SY46	TTC9	24	Q92623	K61 [HEFK(me1)SQG]
pan-Kme1	TPD52	24	P55327	K149 [AGQK(me1)ASA]
pan-Kme1	RPL13	24	P26373	K177 [KNFK(me1)AFA]
pan-Kme1	CHMP2B	24	Q9UQN3	K6 [SLFK(me1)KKT]
pan-Kme1	CHMP2B	24	Q9UQN3	K7 [LFKK(me1)KTV]
pan-Kme1	RPL14	23	P50914	K53 [MPFK(me1)CMQ]
pan-Kme1	GSTP1	23	P09211	K45 [GSLK(me1)ASC]
pan-Kme1	TPD52L2	22	O43399	K90 [GELK(me1)QNL]
both	BTF3	22	P20290-2	K2 [_ _MK(me1)ETI]
pan-Kme1	BTF3	22	P20290	K46 [PQMK(me1)ETI]
pan-Kme1	BTF3	22	P20290	K57 [KLAK(me1)LQA]
pan-Kme1	COMMD3	22	Q9UBI1	K137 [YQIK(me1)TNQ]
pan-Kme1	HIST1H1C	21	P16403	K46 [LITK(me1)AVA]
both	COMMD1	21	Q8N668	K54 [FLAK(me1)MRG]
pan-Kme1	UBE2M	21	P61081	K3 [_MIK(me1)LFS]
pan-Kme1	TRAPPC5	21	Q8IUR0	K9 [TRGK(me1)SAL]
both	PDAP1	21	Q13442	K126 [RYMK(me1)MHL]
both	TDRP	20	Q86YL5	K145 [GSTK(me1)YTS]
pan-Kme1	COPZ1	20	P61923	K172 [EQIK(me1)WSL]
both	HN1L	20	Q9H910	K184 [PGGK(me1)SSI]
pan-Kme1	HN1L	20	Q9H910	K67 [PGGK(me1)GSG]
pan-Kme1	HN1L	20	Q9H910	K90 [PGGK(me1)TSD]
pan-Kme1	ARPC4	20	P59998	K44 [RSSK(me1)ELL]

Tab. 7.10: Detected lysine mono-methylation sites

Antibody	**Gene**	**Mass, kDa**	**Uniprot ID**	**Kme1 site**
pan-Kme1	CNBP	19	P62633	K8 [ECFK(me1)CGR]
pan-Kme1	PCNP	19	Q8WW12	K81 [QTTK(me1)KAS]
pan-Kme1	PCNP	19	Q8WW12	K82 [TTKK(me1)ASA]
SY46	MRPL30	19	Q8TCC3	K140 [TCLK(me1)STG]
pan-Kme1	RPS11	18	P62280	K38 [RYYK(me1)NIG]
pan-Kme1	C19orf43	18	Q9BQ61	K128 [LALK(me1)TGI]
pan-Kme1	PPIA	18	P62937	K49 [FGYK(me1)GSC]
SY46	PPIA	18	P62937	K91 [FILK(me1)HTG]
both	NCBP2	18	P52298	K7 [GLLK(me1)ALR]
pan-Kme1	C5orf49	17	A4QMS7	K61 [CIFK(me1)RRL]
SY46	CALML3	17	P27482	K78 [RKMK(me1)DTD]
pan-Kme1	EDF1	16	O60869	K25 [AKSK(me1)QAI]
pan-Kme1	C1D	16	Q13901	K119 [RFVK(me1)NAL]
pan-Kme1	HN1	16	Q9UK76	K148 [PGGK(me1)SSL]
pan-Kme1	HN1	16	Q9UK76	K8 [TTFK(me1)GVD]
pan-Kme1	SOD1	16	P00441	K10 [CVLK(me1)GDG]
pan-Kme1	POLR1E	16	Q9GZS1-2	K80 [GALK(me1)CNT]
both	HIST1H3A	15	P68431	K80 [QDFK(me1)TDL]
SY46	MPC1L	15	P0DKB6	K10 [LWRK(me1)MRD]
SY46	MPC1L	15	P0DKB6	K18 [FQSK(me1)EFR]
pan-Kme1	SRP14	15	P37108	K107 [KKTK(me1)AAA]
pan-Kme1	PPP1R11	14	O60927	K59 [RSSK(me1)CCC]
pan-Kme1	SNRPD3	14	P62318	K104 [AILK(me1)AQV]
pan-Kme1	TSTD1	13	Q8NFU3	K83 [QMGK(me1)RGL]
both	RPL36A	12	P83881	K53 [GQTK(me1)PIF]
both	HIST1H4A	11	P62805	K21 [RHRK(me1)VLR]
SY46	HIST1H4A	11	P62805	K92 [YALK(me1)RQG]
pan-Kme1	CSTB	11	P04080	K44 [VSFK(me1)SQV]
pan-Kme1	RPS21	9	P63220	K51 [GQFK(me1)TYA]
pan-Kme1	GNG12	8	Q9UBI6	K4 [MSSK(me1)TAS]
pan-Kme1	TMA7	7	Q9Y2S6	K50 [AAGK(me1)GPL]

7.3 Gene ontology analyses of genes corresponding to methylated peptides detected in proteom study

Tab. 7.11: Gene ontology analysis of genes corresponding to methylated peptides in proteom study using AmiGO web service.

GO biological process complete	Fold Enrichment	P value
retrograde vesicle-mediated transport, Golgi to ER	5.48	2.56E-02
regulation of viral genome replication	5.41	2.91E-02
DNA duplex unwinding	5.35	3.31E-02
telomere organization	5.21	4.25E-02
positive regulation of viral life cycle	5.08	2.28E-02
protein localization to endoplasmic reticulum	4.22	3.73E-02
Golgi vesicle transport	3.8	6.55E-07
microtubule-based movement	3.69	2.37E-03
DNA replication	3.42	1.40E-02
DNA conformation change	3.36	3.48E-03
ribosome biogenesis	3.29	2.09E-04
rRNA processing	3.27	5.60E-03
RNA splicing, via transesterification reactions	3.22	4.36E-03
rRNA metabolic process	3.2	8.31E-03
mRNA splicing, via spliceosome	3.14	1.14E-02
establishment of protein localization to membrane	3.12	1.30E-02
ribonucleoprotein complex biogenesis	2.99	2.28E-05
microtubule cytoskeleton organization	2.78	1.57E-02

Tab. 7.12: Gene ontology analysis of genes corresponding to methylated peptides in proteom study using AmiGO web service.

GO molecular function complete	Fold Enrichment	P value
ATP-dependent DNA helicase activity	7.41	4.48E-02
cadherin binding involved in cell-cell adhesion	6.48	7.77E-22
DNA helicase activity	6.28	1.70E-02
protein binding involved in cell-cell adhesion	6.27	3.15E-21
protein binding involved in cell adhesion	6.17	6.70E-21
microtubule motor activity	5.28	1.21E-02
ribonucleoprotein complex binding	4.62	2.01E-02
actin filament binding	4.36	3.84E-03
cell adhesion molecule binding	4.22	5.59E-15
microtubule binding	3.85	1.89E-04
poly(A) RNA binding	3.78	2.47E-34
tubulin binding	3.56	2.20E-05
actin binding	3.48	5.77E-08
GTPase binding	3.27	2.15E-04
RNA binding	3.17	1.73E-32
cytoskeletal protein binding	3.11	3.80E-14

Tab. 7.13: Gene ontology analysis of genes corresponding to methylated peptides in proteom study using AmiGO web service.

GO cellular component complete	Fold Enrichment	P value
cell-cell adherens junction	5.63	1.70E-19
nucleolar part	5.59	2.23E-02
microtubule associated complex	3.99	2.91E-03
ruffle	3.79	5.73E-03
adherens junction	3.78	9.36E-19
anchoring junction	3.73	9.68E-19
cell-cell junction	3.6	2.05E-15
spliceosomal complex	3.6	6.02E-03
coated vesicle	3.53	2.28E-04
postsynaptic specialization	3.5	4.88E-03
nuclear speck	3.39	7.52E-03
excitatory synapse	3.29	1.14E-02
cell-substrate junction	3.18	3.01E-06
focal adhesion	3.15	6.51E-06
cell-substrate adherens junction	3.13	7.87E-06
nucleolus	3	1.39E-13
cell leading edge	2.87	8.52E-04
actin cytoskeleton	2.87	2.62E-05
microtubule	2.86	1.63E-04

7.4 Exemplary MS/MS spectra of methylated peptides of AHNAK-CRUs after biochemical methylation with SMYD2

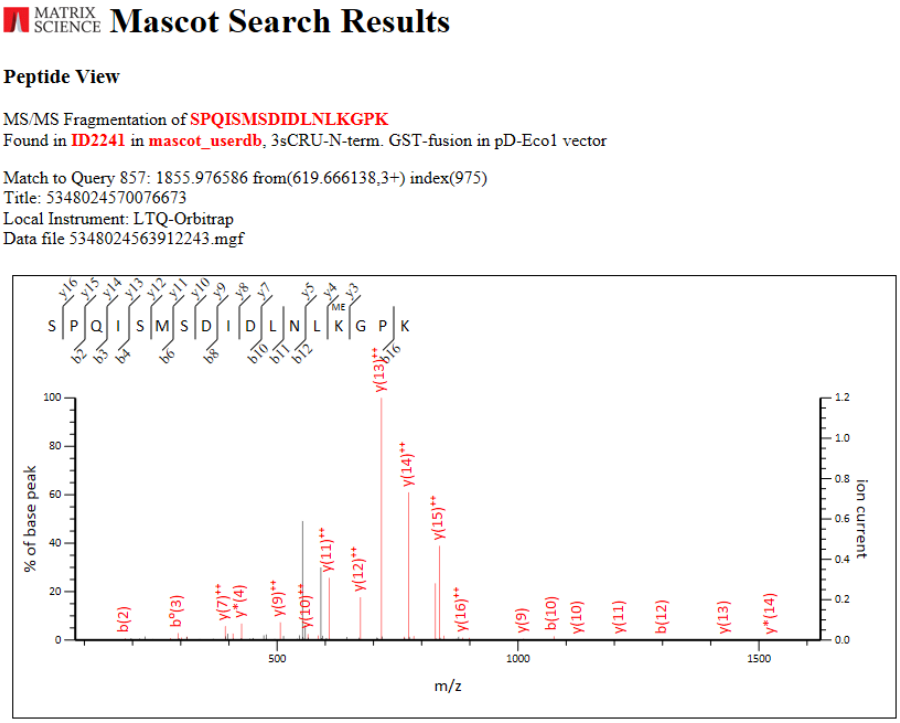


Fig. 7.49: SMYD2 in vitro methylates AHNAK-CRUs at LK*GP motifs.
Exemplary MS/MS spectrum of peptide from AHNAK-3sCRU after in vitro methylation with SMYD2. Monomethylation was detected at lysine within the LK*GP motif. For detection, protein bands from the gel shown in **Fig. 4.36** were cut off and processed for mass spectrometry analysis.

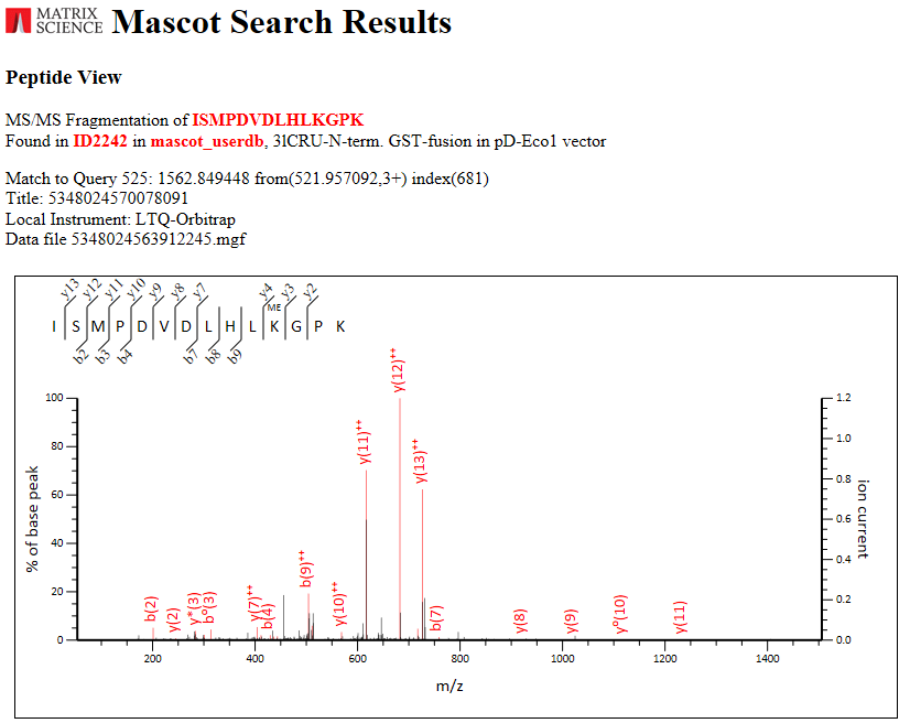


Fig. 7.50: SMYD2 in vitro methylates AHNAK-CRUs at LK*GP motifs.
Exemplary MS/MS spectrum of peptide from AHNAK-3lCRU after in vitro methylation with SMYD2. Monomethylation was detected at lysine within the LK*GP motif. For detection, protein bands from the gel shown in **Fig. 4.36** were cut off and processed for mass spectrometry analysis.

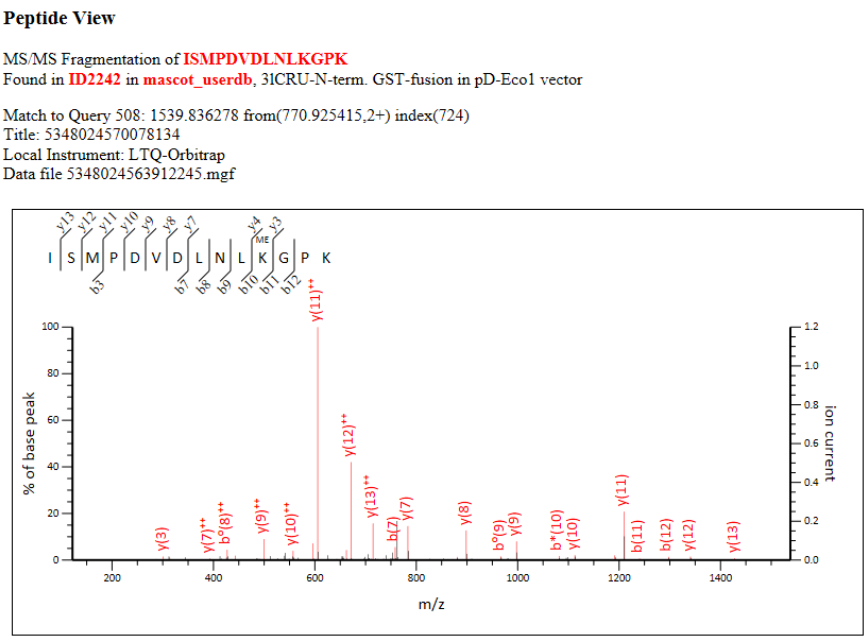


Fig. 7.51: SMYD2 in vitro methylates AHNAK-CRUs at LK*GP motifs.
Exemplary MS/MS spectrum of peptide from AHNAK-3lCRU after in vitro methylation with SMYD2. Monomethylation was detected at lysine within the LK*GP motif. For detection, protein bands from the gel shown in **Fig. 4.36** were cut off and processed for mass spectrometry analysis.

7.5 Localization of SMYD2 after cellular fractionation

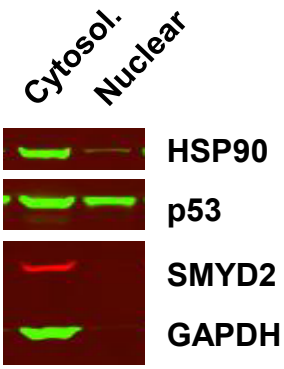


Fig. 7.52: SMYD2 is localized to the cytoplasm.
Parental KYSE150 cells were lysed and fractionated into cytoplasmic and nuclear fractions using a cytoplasmic/nuclear extraction kit (Thermo Fisher Scientific, #78833) according to the manufactures protocol. Samples were analyzed by western blot and probed for SMYD2, p53, GAPDH and HSP90.

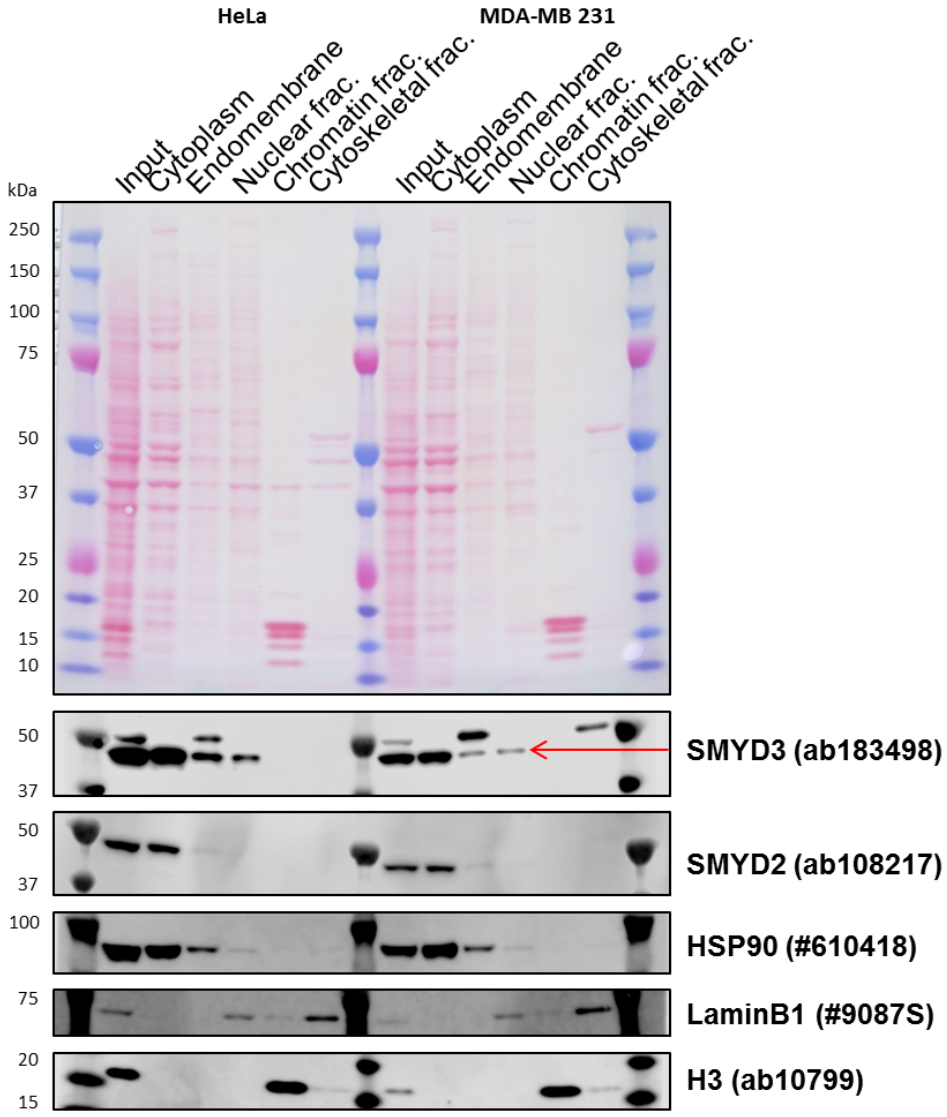


Fig. 7.53: SMYD2 is in the cytoplasmic fraction.
HeLa and MDA-MB231 parental cell pellets were prepared for subcellular fractionation using Subcellular Protein Fractionation Kit for Cultured Cells (Thermo Fisher Scientific, #78840) according to the manual and analyzed by western blot. Input means the whole cell extract before fractionation (in brackets are the antibody cat.#, SMYD2, SMYD3, H3 from abcam, HSP90 from BD Pharmingen, LaminB1 from Cell Signaling Technology). SMYD2 and SMYD3 are in the cytoplasm. HSP90 is mainly in the cytoplasm and to some extend in the membrane fraction. Histone H3 and LaminB1 served as controls for nuclear and chromatin fractions, respectively.